UNIVERSITY OF CALGARY

Fine-Grained Quantum Uncertainty Relations

by

Alireza Poostindouz

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN PHYSICS AND ASTRONOMY

CALGARY, ALBERTA

September, 2016

Abstract

Quantum theory predicts an inherent joint unpredictability for some pairs of measurements. For example, Heisenberg showed that the more precisely the position of a quantum particle is known, the less precisely its momentum can be known and vice versa. Uncertainty relations (URs) are mathematical expressions quantifying the constraints between the output probability distributions of the given sets of measurements. Typically, URs are expressed in terms of uncertainty quantifiers such as entropies. Based on an information-theoretic approach, we discovered a characterization that unifies all uncertainty quantifiers and thus, generalizes a large class of URs into a single framework. We also prove new URs that are fundamentally different from typical URs in that they are fine-grained; i.e. they set restrictions directly on the output probability distributions without using any particular uncertainty quantifiers. We used Majorization theory and other techniques such as matrix analysis to prove our fine-grained uncertainty relations.

Acknowledgements

First and foremost I would like to dedicate this thesis to my loving wife Parisa Zarkeshian for her patience, kind support, and encouragements. Without her I wouldn't be able to finish this program.

Special thanks go towards my supervisor Dr. Gilad Gour for his guidance, and patience over the last two years. I would also like to express my gratitude to my supervisory committee members, Dr. Peter Høyer and Dr. Barry Sanders. I am truly grateful not also for being able to work under supervision of these three outstanding scientists but also for the opportunity to learn a lot about quantum information science from three excellent teachers.

I would like to thank my good friends and many grad students of the Institute for Quantum Science and Technology; particularly Varun Narasimachar, Mark Girard, Majid Shahabi, Mehdi Ahmadi, Hoan Bui Dang, William Johnson, Abdullah Khalid, Borzumehr Toloui Semnani, and Sourabh Kumar. Also, I would like to acknowledge NSERC and the Department of Physics and Astronomy for their financial support. Special mention goes to the nice and kind staff of IQST and the Department of Physics and Astronomy, especially Tracy Korsgaard, Gerri Evans, Nancy Jing Lu, and Lucia Wang.

Last but not least, I am sincerely thankful to my parents and my sweet little sister for all their sacrifices, their continual support and all the things they taught me. There are not enough words in any languages which can let me express my gratitude for my beloved parents.

Table of Contents

List of Figures and Illustrations

List of Symbols, Abbreviations and Nomenclature

Symbol	Definition
DMC	Discrete memoryless channel
POVM	Positive operator-valued measurement
UR	Uncertainty Relation
EUR	Entropic Uncertainty Relation
$H_1(\boldsymbol{p})$	Shannon entropy of probability vector $\boldsymbol{p}$
$H_\alpha(\boldsymbol{p})$	Rényi entropy of probability vector $\boldsymbol{p}$
$\mathcal{X}, \mathcal{Y}, \mathcal{Z}$	Alphabets of random variables X, Y, Z
σ_X^2	Variance of random variable X
$\mathbb{E}\{X\}$	Expected value of random variable X
$T_{y\|x}$	Classical channel with input X and output Y
P_π	Matrix representation of permutation π
$\boldsymbol{q} \prec \boldsymbol{p}$	$\boldsymbol{p}$ majorizes $\boldsymbol{q}$
$\boldsymbol{e} = (1, \cdots, 0)$	Completely certain distribution
$\boldsymbol{u} = (1/d, \cdots, 1/d)$	Uniform distribution
$\boldsymbol{p}^\downarrow$	Sorted version of $\boldsymbol{p}$ in a non-increasing order
$D = (D_{kj})$	Doubly stochastic matrix
$\Phi(\boldsymbol{p})$	Schur-concave function
$\mathcal{H}$	Hilbert space
$\mathsf{dim}(\mathcal{H})$	Dimension of $\mathcal{H}$
$\mathsf{span}(\mathcal{H})$	Span of $\mathcal{H}$
$\mathsf{L}(\mathcal{H})$	Space of all linear operators of $\mathcal{H}$

$\boldsymbol{\lambda}(\rho)$	Vector of non-zero eigenvalues of density operator ρ
Π_m	Projector operator
$\hat{\mathrm{O}}$	Measurement observable operator
$\boldsymbol{\mathcal{A}} = \{\Lambda_j\}_j$	Quantum POVM measurement $\boldsymbol{\mathcal{A}}$
$\boldsymbol{\mathcal{B}} = \{\Gamma_j\}_j$	Quantum POVM measurement $\boldsymbol{\mathcal{B}}$
$\mathrm{Tr}\{\Lambda\}$	Trace of operator Λ
$\mathcal{U}(\boldsymbol{p})$	Single-variable uncertainty quantifier
$\mathcal{J}(\boldsymbol{p}, \boldsymbol{q})$	Joint uncertainty quantifier
$\mathcal{D}$	Set of all doubly stochastic matrices

Chapter 1

Introduction

Quantum theory portrays a spooky picture of the universe. One of the distinctive properties of quantum theory is the inherent unpredictability of quantum measurements. Heisenberg showed that measurements of position and momentum for a single quantum particle are *incompatible*: the more precisely the position of a particle is known, the less precisely its momentum can be known and vice versa [1].

An *uncertainty relation* is a quantitative statement expressing unpredictability constraints associated with a set of quantum measurements. There exist various uncertainty relations, and each of them use diverse mathematical functions to quantify the uncertainty associated with a quantum measurement. For example, in his famous uncertainty relation, Heisenberg uses standard deviation of position and momentum as the uncertainty quantifier, while information theorists, mostly use *entropies* as the preferred uncertainty quantifiers [2].

Different uncertainty relations use not only different uncertainty quantifiers, but also different approaches to express the connection between the uncertainties associated with the measurements. Typically, the collective uncertainty associated with the set of two (or more) measurements is quantified by a mathematical function constructed from a combination of uncertainty quantifiers. For instance, Heisenberg uses the *multiplication* of the two standard deviations to describe his uncertainty relation, whereas in other types of uncertainty relations, the *addition* of two entropies are being used. We call the mathematical functions that quantify the collective uncertainty associated with two or more measurements *joint uncertainty quantifiers* or *quanti-*

fiers of joint uncertainty. Most uncertainty relations are formulated in terms of joint uncertainty quantifiers.

In this thesis, our primary goal is to use information-theoretic principles to identify the "common essence" of all uncertainty quantifiers, thereby defining notions of uncertainty and joint uncertainty independent of any quantifiers. Our formalism unifies all existing quantifiers of uncertainty and joint uncertainty. But, the important result of this work is that it unifies a large class of quantum uncertainty relations into a single framework. Also, to illustrate the utility of our formalism, we derive a new quantum uncertainty relation based on an unexplored joint uncertainty quantifier (see [3]). Ultimately, we use our formalism to prove a new class of fine-grained uncertainty relations (see section 4.3).

In this work we only focus on the 'quantum uncertainty of preparation'. The following example clarifies the concept.

1.1 Example: Quantum Uncertainty of Preparation

How can two quantum measurements be uncertain at the same time? Consider the following simple example: "Measuring a *qubit* with respect to two different bases".

A *qubit* (quantum bit) is a two-level quantum system, e.g. the spin of an electron, or the polarization of a single photon. The state of a qubit is described by a norm-one 2-dimensional complex vector $|\psi\rangle$ in $\mathbb{C}^2$. Assume that the description of an arbitrary qubit $|\psi\rangle$ in the canonical basis $\{|0\rangle, |1\rangle\}$ is:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{1.1}$$

where $\alpha, \beta \in \mathbb{C}$ with $|\alpha|^2 + |\beta|^2 = 1$. For simplicity, assume that α and β are real. So, we can visualize our qubits with unit vectors in a $2D$ plane.

We model quantum measurement with respect to $\{|0\rangle, |1\rangle\}$ by a device that inputs a quantum state $|\psi\rangle$ and outputs an scalar value: either 0 or 1. Imagine that we can prepare the input qubit as

$$|\psi\rangle = \cos(\theta)|0\rangle + \sin(\theta)|1\rangle. \tag{1.2}$$

Then, the measurement outputs 0 and 1 with probabilities p_0 and p_1, respectively:

$$p_0 = \cos^2(\theta), \tag{1.3}$$

$$p_1 = \sin^2(\theta). \tag{1.4}$$

Consequently, for some arbitrary preparation of $|\psi\rangle$ (for an arbitrary θ), we will not necessarily be able to predict the outcome the measurement with full certainty. We can only calculate the output probability distribution of the measurement, (p_0, p_1). This kind of randomness in the output of a quantum measurement can be removed. For instance, we can prepare the state in $|\psi\rangle = |0\rangle$ (which corresponds to $\theta = 0$) and thus obtain $(p_0 = 1,\ p_1 = 0)$, i.e. a fully certain (predictable) outcome. We can also prepare the state in $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ (which corresponds to $\theta = \frac{\pi}{4}$) and thus obtain $(p_0 = 0.5,\ p_1 = 0.5)$, that is the fully uncertain (unpredictable) outcome.

The canonical basis is not though the only available choice we have for our measurement basis. In fact, we can perform measurements with respect to any other bases. For example, consider the measurement with respect to the basis $\{|+\rangle, |-\rangle\}$, where $|+\rangle \equiv (|0\rangle + |1\rangle)/\sqrt{2}$ and $|-\rangle \equiv (|0\rangle - |1\rangle)/\sqrt{2}$. As the description of the same qubit (see (1.2)), in $\{|+\rangle, |-\rangle\}$ is

$$|\psi\rangle = \sin\left(\theta + \frac{\pi}{4}\right)|+\rangle + \cos\left(\theta + \frac{\pi}{4}\right)|-\rangle, \tag{1.5}$$

the resulting output probabilities of measuring this qubit with respect to the new

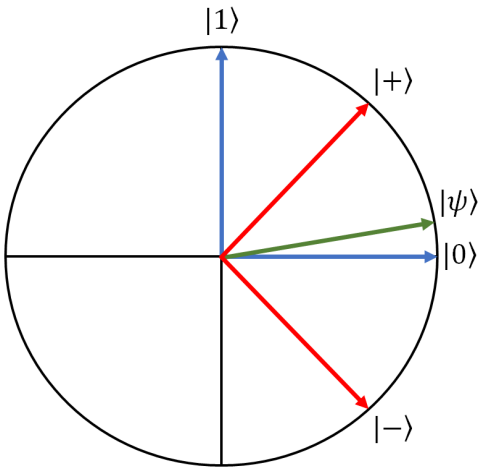


Figure 1.1: Visualization of quantum states in a $2D$ plane.

basis is:

$$q_+ = \sin^2\left(\theta + \frac{\pi}{4}\right), \tag{1.6}$$

$$q_- = \cos^2\left(\theta + \frac{\pi}{4}\right); \tag{1.7}$$

where q_+ and q_- are the probabilities of obtaining $|+\rangle$ and $|-\rangle$, respectively. See Fig 1.1 for a visualization of $|\psi\rangle$ and the two measurements bases.

Now consider the relationship between the output probability distributions associated with the set of two quantum measurements, (p_0, p_1), and (q_+, q_-). There exists a constraint on the values that (p_0, p_1) and (q_+, q_-) can take: we can never prepare our qubit in such a way that makes (p_0, p_1) and (q_+, q_-) fully certain together (see Fig. 1.2). For example, if the qubit is prepared in the state $|\psi\rangle = |0\rangle$, the first output distribution is completely certain (or totally predictable), i.e. $(p_0 = 1, p_1 = 0)$, whereas the second measurement is completely uncertain (or totally unpredictable); i.e. $(q_+ = 1/2, q_- = 1/2)$.

We can quantify these quantum uncertainty constraints with many different ways. For instance, we can prove that for a given p_0 the value of q_+ resulted from the very

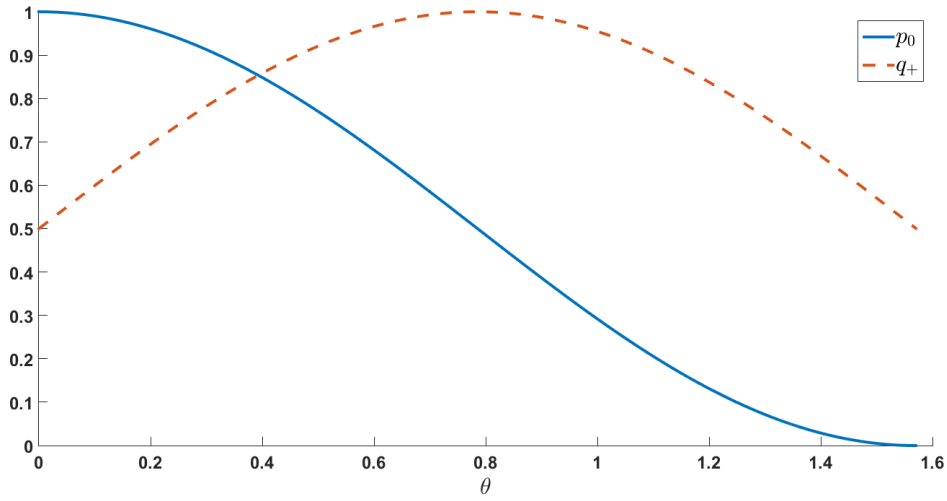


Figure 1.2: Constraints on the output probability distributions: p_0 and q_+ cannot be 0 or 1 for a same θ. For preparations (θ's) for which p_0 is more certain (closer to 0 or 1), q_+ becomes more uncertain (closer to 1/2) and vice versa.

same preparation is always either of the following: $q_+ = \frac{1}{2} \pm \sqrt{p_0(1-p_0)}$. The important fact is that if the two given probability distributions are simulated by two classical random processes they will never be correlated in the exact same way that (p_0, p_1) and (q_+, q_-) are correlated here. This was one simple example of the quantum uncertainty of preparation.

In summary, quantum mechanics predicts some specific uncertainty constraints for a given set of quantum measurements. If these measurements are incompatible (non-commuting), then according to the laws of quantum mechanics any state preparation that makes either of the measurements more certain makes the other more uncertain.

In this thesis we only study the uncertainty relations and uncertainty constraints of the output probability distributions resulted from a given set of quantum measurements in the preparational setting. Our results, theorems and uncertainty relations are not necessarily hold for other probability distributions simulated by any classical stochastic processes.

1.2 Background and Motivation

Historically, the first and thus far the most famous statement about the inherent uncertainty about the outputs of a pair of quantum measurements is the Heisenberg's uncertainty principle. He showed that position and momentum measurements of a single quantum particle are jointly uncertain (meaning that at least for some preparations, one of the measurements is uncertain). This quantum uncertainty can be quantitatively stated via the following uncertainty relation (UR)

$$\Delta x \Delta p \geq \frac{\hbar}{2}, \tag{1.8}$$

where Δx and Δp are respectively the standard deviations of position and momentum measurements, and $\hbar$ is the reduced Planck constant [1].

There exist many quantum uncertainty relations, each capturing one aspect of the quantum uncertainty principle. Some of these URs use standard deviation as their uncertainty quantifier, whereas others are formulated in terms of entropies. One well-known example of an entropic uncertainty relation (EUR) is the relation proved by Maassen and Uffink [4]. Let $H_1(A) = -\sum_x p_x \log p_x$ denote the Shannon entropy associated with measurement A with respect to $\{|a_x\rangle\}$, where p_x is the probability of obtaining $|a_x\rangle$ at the output of the measurement. Maassen and Uffink showed that for two quantum measurements A and B with respect to two orthonormal bases $\{|a_x\rangle\}$ and $\{|b_y\rangle\}$

$$H_1(A) + H_1(B) \geq -2 \log \eta(A, B), \tag{1.9}$$

where $H_1(A)$ and $H_1(B)$ are respectively the Shannon entropies associated with output probability distributions of A and B, and $\eta(A, B)$ is a positive constant depending only on the two quantum measurements A and B.

In this thesis, we introduce a general framework that unifies a large class of uncertainty relations. The importance of gaining a better insight about a fundamental feature of the quantum theory is a strong motivation for our efforts. In addition, uncertainty relations, especially EURs, have a lot of applications in many research areas including quantum information theory (see for example [5–7]). For a comprehensive list of works and applications of different URs we point to these excellent recent surveys [8, 9].

Despite the growing trend for optimizing the lower bounds of different EURs (e.g. [10–13]), the aim of this thesis is to study the concepts of uncertainty, joint uncertainty, and quantum uncertainty relations from an information-theoretic perspective. Majorization theory and quantum mechanics form the foundation of this thesis. A quick review of the notions that are vital for understanding the main results of this work is presented in chapter 2. We explain our formalism and our axiomatic definitions of valid uncertainty and joint uncertainty quantifiers in chapter 3. Employing the techniques developed in chapter 3, we prove results about quantum mechanical uncertainty relations in chapter 4. We conclude this thesis with our conclusions and a discussion about possible future work in chapter 5.

1.3 Clarifications and Summary of Results

In the literature, the uncertainty of a variable has almost always been discussed in terms of functions that quantify "the amount of uncertainty". The famous examples of such quantifiers are the Shannon entropy and its extended family of Rényi entropies, geometric norm-based quantifiers such as the quadratic variance, etc. In most cases, there is a clear operational meaning for such quantifiers, rendering them well-suited to their various applications. Similarly, when two or more experiments with random

outputs are under study, quantifiers of joint uncertainty have been constructed by two usual approaches: either by combining single-variable uncertainty quantifiers mathematically, or by considering operational combined tasks that involve all the experiments of interest.

In the present work, we extract the common thread beneath the operational descriptions of all such (single or joint) uncertainty quantifiers. Our argument is based on some simple operational axioms that are independent of the quantifier used to quantify uncertainty. Therefore, our approach gives a more intrinsic look at the concept of uncertainty.

The axioms we will introduce are motivated by information-theoretic principles and are intended to be as objective as possible. Considering the inherent challenges in such a requirement, we restrict the generality of our treatment in the following ways. First, we only study the notions of uncertainty applicable to classical random variables. In particular, this class of variables includes the classical outcomes of quantum-mechanical measurements. Second, we restrict our work to discrete random variables; in fact, we consider only random variables with finite outcome space. We make some tentative suggestions about the treatment of discrete and continuous infinite-dimensional cases, but leave the actual extension for future work. Finally, to compare the uncertainties of different variables, we will require the compared variables to represent the same type of physical quantity. For example, comparing the uncertainties of two different length variables will be possible within our formalism, but comparison between a length uncertainty and a mass uncertainty is not allowed.

The crux of this thesis are the following axioms:

1. Our knowledge of a variable cannot increase under any processing without addition of new information about the variable.

2. The uncertainty in a variable representing a physical observable is invariant under the symmetry transformations of the observable.

3. The joint uncertainty of several variables is a valid concept even without an underlying operational description that combines those variables.

The first two axioms are inspired by earlier approaches [14–16] to quantifier-independent notions of uncertainty, wherein the connection between uncertainty and a mathematical concept called majorization was utilized. Majorization is a hierarchy among probability distributions which is induced by the action of a class of transformations called doubly stochastic maps (see [17]).

In chapter 3, we find a mathematical characterization of mechanisms that can increase a variable's uncertainty. In fact, we gain an understanding of why, and to what extent, majorization plays a role in characterizing uncertainty. For variables with unrestricted symmetries we find that uncertainty-increasing mechanisms are associated with the set of *all* doubly stochastic maps. This leads us to the emergence of majorization as the relation determining uncertainty. Thus in this case, a function that quantifies uncertainty must never decrease under the action of any doubly stochastic map on the probability distribution. On the other hand, for variables with restricted symmetries, only certain sub-classes of doubly stochastic maps feature. The resulting hierarchy is different from majorization, and therefore a quantifier of uncertainty is only required to be non-decreasing under the action of the restricted classes of doubly stochastic maps. This opens up a bigger variety of functions that can serve as uncertainty quantifiers for variables with restricted symmetries.

Another element of novelty in our work lies in our third axiom, which concerns joint uncertainty. Operationally, we can rephrase this axiom in terms of experiments with random outputs. For example, suppose that we are interested in quantifying

the joint uncertainty about the set of random outcomes of several quantum measurements. One approach would be to construct new experiments that combine the original experiments in some way. For example, consider the following two combined experiments each of which constructed from a given set of experiments:

(a) All the original experiments are performed independently.

(b) All the experiments' apparatuses are set up, but only one of the experiments is chosen at random and performed.

The uncertainty in the outcome of such a combined experiment would quantify the joint uncertainty of the constituent experiments. However, we note that there are different ways to combine experiments, each of which capturing different aspects of the joint uncertainty.

In chapter 4, we argue that the richness of joint uncertainty is not captured even by considering all such combined experiments. The most general notion of joint uncertainty is devoid of the particulars of such combinations, and allows all the component experiments to be, in principle, counterfactual. To illustrate this, we consider an extensively-studied type of quantum uncertainty relations: the so-called preparational uncertainty relations. For ease of explanation, let us consider a two-measurement[1] preparational uncertainty relation, which has the generic form

$$\mathcal{J}\left(\boldsymbol{p}(\rho), \boldsymbol{q}(\rho)\right) \geq c, \tag{1.10}$$

where $\mathcal{J}$ is a quantifier of the joint uncertainty of two variables, and $\boldsymbol{p}(\rho)$ and $\boldsymbol{q}(\rho)$ are the expected outcome probability distributions of a pair of measurements performed on a quantum state represented by the density operator ρ. We show that most existing

[1]Our arguments can be extended to more than two measurements.

preparational uncertainty relations can be subjected to one of the specific operational interpretations (a) and (b) mentioned above. With some efforts, these restrictions can be removed. In fact, we construct joint uncertainty quantifiers that cannot be interpreted either way. In addition, we derive a quantum uncertainty relation based on one such quantifier, which is different from all the ones discovered in the past. The main purpose of deriving this new relation is to demonstrate that our axioms open up the possibility of finding infinitely new joint uncertainty quantifiers, and thus new quantum uncertainty relations.

All existing two-measurement quantum uncertainty relations have the generic form of (1.10); however, the essence of all these relations is hidden behind the constraints they predict for the output probability distributions of the set of relevant measurements. On this account, no single inequality (which sets a positive lower bound on some function $\mathcal{J}(\boldsymbol{p}(\rho), \boldsymbol{q}(\rho))$) can capture all aspects of *the unpredictability* about the outcomes of a pair of quantum measurements[2].

Fine-grained uncertainty relations are mathematical assertions that explain the unpredictability constraints directly in terms of the output probability distributions of the given set of quantum measurements. We call them also "quantifier-free URs" as they can imply a large class of URs and they are formulated independent of any single uncertainty (or joint uncertainty) quantifier. Motivated by the earlier efforts [14, 15], we therefore tried to find *the universal* quantifier-free formulation of all uncertainty relations. More precisely, we aimed to find a fixed pair of vectors $(\boldsymbol{s}, \boldsymbol{t})$ (independent of ρ) such that $\mathcal{J}(\boldsymbol{s}, \boldsymbol{t})$ provides a nontrivial bound [like the c of Eq. (1.10)] for a whole class of quantifiers, $\mathcal{J} \in \mathsf{J}$. We find that no such universal relation exists if J includes all possible quantifiers.

[2] All of our results explained to this point are also covered in [3].

We also argue that a so-called universal majorization relation found in [15] is indeed not universal, because it only covers a restricted class of joint uncertainty quantifiers: the class of all uncertainty quantifiers adhere to type (a) of combined experiments where all the quantum measurements are performed independently on the same quantum state. Similarly, we find a new class of fine-grained majorization uncertainty relations for another restricted set of joint uncertainty quantifiers: the class of all uncertainty quantifiers adhere to type (b) of combined experiments where each time only one quantum measurement is picked at random and will be performed on the input state[3].

Even though the focus of this work is on preparational uncertainty relations in quantum mechanics, in principle the introduced notions can be applied to any situation where probability-based uncertainty quantifiers of classical variables are relevant.

1.3.1 List of Contributions

To conclude this chapter, we list below the contributions of this thesis. (Contributions 1-6 are also covered and discussed in [3].)

1. Based on simple information-theoretic axioms, we give a mathematical characterization that unifies all single-variable uncertainty quantifiers into a single framework. (See section 3.3 and claims 3.1 and 3.2. Also see definition 2 on page 5 of [3].)

2. Similarly, we find a mathematical characterization of all joint uncertainty quantifiers such as, $\Delta x \Delta p$ and $H_1(A) + H_1(B)$. (Discussed in section 3.4 and definition 3.2 and also section 2 of [3].)

[3]The results regarding quantifier-free relations are not covered in [3], for the full discussion please read section 4.3.

3. We argue that joint uncertainty is a concept devoid of all single-variable uncertainty quantifiers and any underlying operational descriptions. (See the examples given in section 3.4, which can also be found on page 6 of [3].)

4. We provide a simple joint uncertainty quantifier (defined in equation (3.16)) as an example to support our previous claim. (Also see page 8 of [3].)

5. We show that all typical URs have the following generic form:

$$\mathcal{J}\left(\boldsymbol{p}(\rho), \boldsymbol{q}(\rho)\right) \geq c, \tag{1.11}$$

   where $\mathcal{J}$ is a quantifier of the joint uncertainty of two variables, and $\boldsymbol{p}(\rho)$ and $\boldsymbol{q}(\rho)$ are the expected outcome probability distributions of a pair of quantum measurements performed on the same quantum state ρ. (Read the discussion in section 4.2 or pages 7 and 8 of [3].)

6. Our formalism enables us to find new URs. To show this, we prove a new unexplored UR in theorem 4.1. (Read section 4.2.1 or see the proof of equation (10) of [3].)

7. We also show that there exists no *universal* set of distributions $(\boldsymbol{s}, \boldsymbol{t})$ such that $\mathcal{J}(\boldsymbol{s}, \boldsymbol{t})$ gives a non-trivial lower-bound for any given arbitrary joint uncertainty quantifier $\mathcal{J}$. (Go to the discussion in section 4.3)

8. Nonetheless, we prove a fine-grained UR (theorem 4.2) that gives us a pair of distributions $(\boldsymbol{s}, \boldsymbol{t})$ such that

$$\mathcal{J}\left(\boldsymbol{p}(\rho), \boldsymbol{q}(\rho)\right) \geq \mathcal{J}(\boldsymbol{s}, \boldsymbol{t}), \quad \forall \mathcal{J} \in \mathsf{J}_2, \tag{1.12}$$

   where J_2 is a subclass of all joint uncertainty quantifiers. (For a complete discussion read the proofs to theorem 4.2 and lemma 4.1.)

Chapter 2

Background

Quantum mechanics and Majorization theory make the cornerstone of this thesis. This chapter gives a brief review of notions and main theorems, which will be used in the rest of this thesis. There are many great books on quantum mechanics and quantum information; among these we point out the books by Wilde [18] and Nielsen and Chuang [19]. The book by Marshall, et al [17] is a useful reference on the theory of majorization.

2.1 Theory of Majorization

In this section, we review the main concepts and theorems of majorization theory that are used in the next chapters. We start by introducing the notations that are used in this thesis for discrete random variables, probability distributions, and discrete classical channels.

2.1.1 Discrete Random Variables

Random variables model random systems. A random variable $X : \Omega \mapsto \mathcal{X}$ is a function from the probability space Ω to some set $\mathcal{X}$. Denote by $p_x = \Pr\{X = x\}$ the probability of each realization $x \in \mathcal{X}$. The probability mass function associated

with a discrete random variable is a column vector

$$\boldsymbol{p}_X = \begin{bmatrix} p_1 \\ \vdots \\ p_{|\mathcal{X}|} \end{bmatrix},$$

I sometimes denote it by $\boldsymbol{p}_X = (p_1, \cdots, p_{|\mathcal{X}|})$, or $\boldsymbol{p}_X = (p_x)$. Note that these should be understood as notations for a column vector, not for a row vector.

2.1.2 Discrete Memoryless Channels

In the context of digital communication systems, a discrete memoryless channel (DMC) is defined as:

Definition 2.1. A *discrete channel* is a system with some input alphabet $\mathcal{X}$, some output alphabet $\mathcal{Y}$ and a probability transition matrix $T_{y|x} = \Pr\{Y = y|X = x\}$, which represents the conditional probability of occurring y at the output if x was sent through the channel. If T is independent of the previous inputs and outputs of the channel, the channel is said to be *memoryless*.

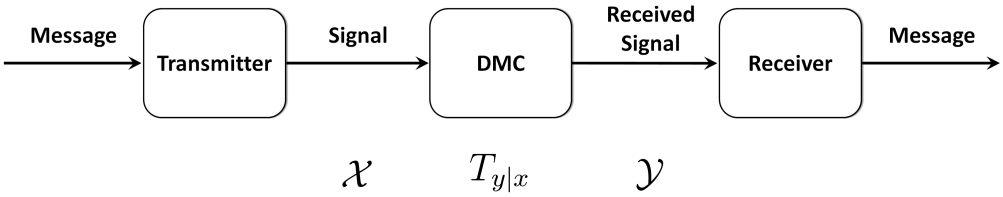


Figure 2.1: Diagram of a digital communication system with a discrete memoryless channel.

Example 2.1. Consider the following binary channel. The input alphabet is $\mathcal{X} = \{0, 1\}$, and the output alphabet is $\mathcal{Y} = \{0, e, 1\}$. For this channel, there are two kinds of probable errors to each input bit, namely erasure and flip bit. Erasure errors

happen when an input x is missed at the receiver. Flip bit errors occur when a "0" in the input transfers to a "1" at the output, or vice versa.

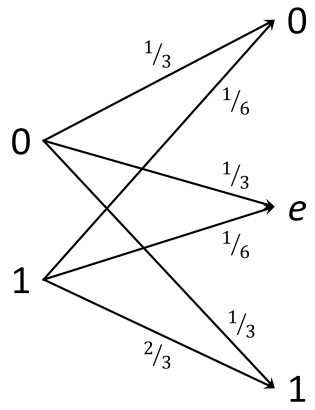


Figure 2.2: An example of a discrete channel.

Based on Fig. 2.2, the transition matrix of this particular channel is the following column stochastic matrix:

$$T = \begin{bmatrix} \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

Note that if the input probability vector is $\boldsymbol{p}_X$, the channel output probability vector is equal to: $\boldsymbol{q}_Y = T\boldsymbol{p}_X$, where T is the column-stochastic matrix representing the channel with $\sum_y T_{y|x} = 1$.

2.1.3 Majorization

Majorization is a relation defined among real vectors. It is a useful mathematical tool to compare between the non-uniformity of two different probability vectors. Let $\boldsymbol{p}^{\downarrow}$ denote the sorted version of $\boldsymbol{p}$ in a non-increasing order, meaning that $p_1^{\downarrow} \geq p_2^{\downarrow} \geq \cdots \geq p_d^{\downarrow}$.

Definition 2.2. The d-dimensional real vector $\boldsymbol{p}$ majorizes the d-dimensional real vector $\boldsymbol{q}$, or in short $\boldsymbol{q} \prec \boldsymbol{p}$ (also can be read as $\boldsymbol{q}$ is majorized by $\boldsymbol{p}$), if and only if

$$\sum_{j=1}^{k} q_j^{\downarrow} \leq \sum_{j=1}^{k} p_j^{\downarrow}, \quad \forall k \in \{1, \cdots, d-1\}, \tag{2.1}$$

and

$$\sum_{j=1}^{d} q_j^{\downarrow} = \sum_{j=1}^{d} p_j^{\downarrow}. \tag{2.2}$$

Obviously, the last equality condition will be satisfied automatically for all probability distribution vectors. Note that the following relations are also equivalent to $\boldsymbol{q} \prec \boldsymbol{p}$, where $\boldsymbol{p}^{\uparrow}$ denotes the sorted version of $\boldsymbol{p}$ in a non-decreasing order:

$$\sum_{j=1}^{k} q_j^{\uparrow} \geq \sum_{j=1}^{k} p_j^{\uparrow}, \quad \forall k \in \{1, \cdots, d-1\}, \tag{2.3}$$

and

$$\sum_{j=1}^{d} q_j^{\uparrow} = \sum_{j=1}^{d} p_j^{\uparrow}. \tag{2.4}$$

Note that it is easy to check whether a vector majorizes the other or not, simply by manipulations of the two vector elements.

Majorization can be regarded as a mathematical tool to compare the non-uniformity between different probability distribution vectors. The intuition to this notion of disorderedness can be understood by the following theorem [17].

Theorem 2.1. *The d-dimensional real vector $\boldsymbol{p}$* majorizes *the d-dimensional real vector $\boldsymbol{q}$, or in short $\boldsymbol{q} \prec \boldsymbol{p}$, if and only if there exist a set of d-dimensional permutation matrices $\{P_\pi\}_\pi$ and a probability vector $\boldsymbol{t} = (t_\pi)$ such that*

$$\boldsymbol{q} = \sum_{\pi} t_\pi P_\pi \boldsymbol{p}. \tag{2.5}$$

Accordingly, "$\boldsymbol{p}$ majorizes $\boldsymbol{q}$" (or $\boldsymbol{q} \prec \boldsymbol{p}$) means that $\boldsymbol{q}$ is more disordered than $\boldsymbol{p}$, since we can generate $\boldsymbol{q}$ merely by randomly permuting $\boldsymbol{p}$ and then averaging over all the permuted versions of $\boldsymbol{p}$. For example, in the 2-dimensional space, we can generate any 2-dimensional probability vector from random permutations of $\boldsymbol{e}_2 = (1, 0)$, or equivalently $\boldsymbol{p} \prec \boldsymbol{e}_2, \forall \boldsymbol{p} \in \mathbb{R}^2$. More precisely, for any 2-dimensional probability vector $\boldsymbol{p} = (p, 1-p)$ we have

$$\boldsymbol{p} = pP_0\boldsymbol{e}_2 + (1-p)P_1\boldsymbol{e}_2,$$

with

$$P_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad P_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

In general, all d-dimensional probability vectors are bounded between the 'completely certain' and 'completely uncertain' distributions $\boldsymbol{e} = (1, 0, \cdots, 0)$ and $\boldsymbol{u} = (1/d, \cdots, 1/d)$, meaning that

$$\boldsymbol{u} \prec \boldsymbol{p} \prec \boldsymbol{e}, \quad \forall \boldsymbol{p} \in \mathbb{R}^d. \tag{2.6}$$

In other words, the fully certain distribution $\boldsymbol{e}$ and all its permuted versions are at the top the majorization hierarchy, meaning that they majorize all other probability vectors of the same dimension. Also, the uniform distribution $\boldsymbol{u}$ and all its permuted versions are at the bottom of majorization hierarchy, which means they can be majorized by all the other distributions.

However, it is important to notice that majorization expresses a *partial ordering* on all probability vectors. Unlike a total order, for which the comparison of any two elements is possible, there exist some cases that neither $\boldsymbol{q} \prec \boldsymbol{p}$ nor $\boldsymbol{p} \prec \boldsymbol{q}$. Two vectors are *incomparable* if neither of them can be generated by different permutations of the other, i.e. $\boldsymbol{q} \not\prec \boldsymbol{p}$ and $\boldsymbol{p} \not\prec \boldsymbol{q}$. For instance, the following vectors are incomparable: $\boldsymbol{p} = (0.5, 0.25, 0.25)$, and $\boldsymbol{q} = (0.4, 0.4, 0.2)$ [17].

2.1.4 Doubly Stochastic Matrices

There is a close connection between doubly stochastic matrices and the theory of majorization. In this section, only basic definitions and some of the main theorems, which are used in the next chapters, will be reviewed. Doubly Stochastic matrices are square matrices in which the entries in every column and every row adds up to one. Therefore, $D = (D_{kj})$ is *doubly stochastic* if $\sum_{j=1}^{d} D_{kj} = \sum_{k=1}^{d} D_{kj} = 1$ and $D_{kj} \geq 0$.

One can view doubly stochastic matrices as square probability transition matrices with the uniform distribution as their stationary state. More precisely, any classical channel D with the same number of inputs and outputs d is doubly stochastic if

$$\boldsymbol{u} = D\boldsymbol{u},$$

where $\boldsymbol{u} = (1/d, \cdots, 1/d)$ is the uniform distribution.

There are two important theorems about doubly stochastic maps. The first one [20] is about characterization of majorization based on doubly stochastic maps. The second one states that the set of $d \times d$ doubly stochastic matrices is identical to the convex hull of all d-dimensional permutation matrices [21]. Here, we state these theorems:

Theorem 2.2 (Hardy, et al. [20])**.** *A d-dimensional real vector $\boldsymbol{p}$ majorizes a d-dimensional real vector $\boldsymbol{q}$ if and only if there exists a $d \times d$ doubly stochastic matrix D such that*

$$\boldsymbol{q} = D\boldsymbol{p}. \tag{2.7}$$

Theorem 2.3 (Brikhoff [21])**.** *Any $d \times d$ doubly stochastic matrix D can be expressed as*

$$D = \sum_{\pi} t_{\pi} P_{\pi}, \tag{2.8}$$

where $\boldsymbol{t} = (t_\pi)$ *is some probability distribution and* $\{P_\pi\}_\pi$ *is the set of all* d*-dimensional permutation matrices.*

Let $\mathcal{D}$ denote the set of all doubly stochastic matrices. The Birkhoff's theorem shows that $\mathcal{D}$ is a convex set with its extreme points being all the permutation matrices P_π.

2.1.5 Schur-Concave Functions

Schur-concave functions are functions that preserve the majorization relation.

Definition 2.3. A function $\Phi : \mathbb{R}^d \mapsto \mathbb{R}$ is *schur-concave* if it is monotonic under majorization, i.e. for two given real vectors $\boldsymbol{p}$ and $\boldsymbol{q}$

$$\boldsymbol{q} \prec \boldsymbol{p} \Rightarrow \Phi(\boldsymbol{p}) \leq \Phi(\boldsymbol{q}). \tag{2.9}$$

Note that if a function is concave, and symmetric under all permutations, it is Schur-concave. It is actually easy to see why. Assume Φ is concave and symmetric under all permutations, i.e.

$$\Phi(P_\pi \boldsymbol{p}) = \Phi(\boldsymbol{p}).$$

Now, if $\boldsymbol{q} \prec \boldsymbol{p}$, then

$$\begin{aligned}
\Phi(\boldsymbol{q}) &= \Phi\left(\sum_\pi t_\pi P_\pi \boldsymbol{p}\right) \\
\Phi(\boldsymbol{q}) &\geq \sum_\pi t_\pi \Phi(P_\pi \boldsymbol{p}) \\
&= \Phi(\boldsymbol{p}),
\end{aligned}$$

where the inequality is due to concavity of Φ and the equality comes from the fact that $\sum t_\pi = 1$. Therefore, any concave function that is symmetric under all permutations is Schur-concave; however, the opposite implication is not true.

Shannon entropy

$$H_1(\boldsymbol{p}) = -\sum_x p_x \log p_x \ , \tag{2.10}$$

and Rényi entropies

$$H_\alpha(\boldsymbol{p}) = \frac{1}{1-\alpha} \log \left(\sum_x p_x^\alpha \right), \tag{2.11}$$

are some important examples of Schur-concave functions. The following theorem provides a method to check Schur-concavity of a given function.

Theorem 2.4 ([17]). *Consider a differentiable function* $\Phi : \mathbb{R}^d \mapsto \mathbb{R}$. Φ *is schur-concave if and only if it satisfies the following conditions:*

1. Φ *is invariant under all permutations, that is*

$$\Phi(P_\pi \boldsymbol{p}) = \Phi(\boldsymbol{p}) \quad \forall P_\pi. \tag{2.12}$$

2. *For all* $\boldsymbol{p} \in \mathbb{R}^d$ *and any indices* k *and* j*,*

$$(p_k - p_j) \left(\frac{\partial \Phi}{\partial p_k} - \frac{\partial \Phi}{\partial p_j} \right) \leq 0. \tag{2.13}$$

To summarize, given two probability distribution vectors $\boldsymbol{p}$ and $\boldsymbol{q}$, all the following statements are equivalent [17]:

1. $\boldsymbol{q} \prec \boldsymbol{p}$.
2. $\boldsymbol{q} = \sum_\pi t_\pi P_\pi \boldsymbol{p}$, for some probability distribution vector $\boldsymbol{t}$ and a set of permutation matrices $\{P_\pi\}$.
3. $\boldsymbol{q} = D\boldsymbol{p}$, for some doubly stochastic matrix D.
4. $\Phi(\boldsymbol{p}) \leq \Phi(\boldsymbol{q})$ for all Schur-concave functions Φ.

2.2 Quantum Mechanics - Quick Review

A simple example of a classical information source is a random bit generator, i.e. a device that outputs zeros and ones at random. For example, consider a device that outputs a random square wave signal of 5 volts amplitude with some clock frequency. We can modulate the classical bits by the following mapping:

$$5 \text{ volts} \mapsto \text{“1”},$$

$$0 \text{ volts} \mapsto \text{“0”}.$$

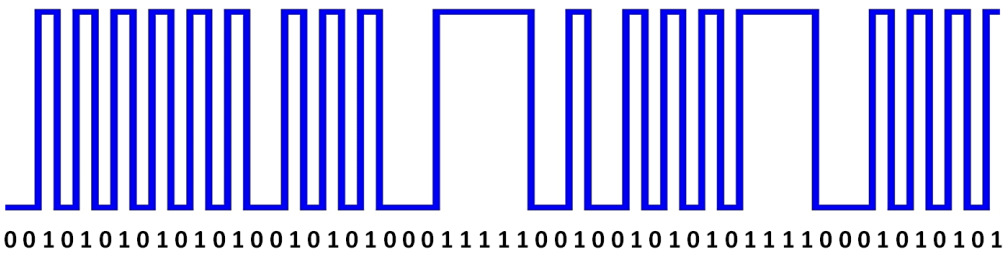


Figure 2.3: An example of a bit stream output of a random bit generator.

In quantum mechanics, states of quantum systems are described by vectors in a Hilbert space $\mathcal{H}$ over the complex field. In the Dirac notation, we denote a column vector associated with the state of a quantum system by ‘ket’ notation, $|\psi\rangle$.

A qubit is a two-dimensional normalized complex vector. It describes the preparation of a two-level system. For instance, the spin of an electron, a two-energy-level quantum system (such as the ground state and exited state), etc.

Now assume we measure a qubit with respect to an orthonormal basis $\{|0\rangle, |1\rangle\}$. One possible outcome of the measurement is $|0\rangle$ which we modulate with the classical bit “0”, and the other possible outcome of the measurement is $|1\rangle$ which we modulate with the classical bit “1”. Thus, we have modeled a classical random bit generator

using a quantum system with the following mapping:

$$|1\rangle \mapsto \text{“1”},$$

$$|0\rangle \mapsto \text{“0”}.$$

However, according to quantum mechanics, $|0\rangle$ and $|1\rangle$ are not the only possibilities for the state of the qubit before the measurement. The state of a qubit can be in any *superposition* of $|0\rangle$ and $|1\rangle$, meaning that the describing state vector of the system is some $|\psi\rangle \in \mathbb{C}^2$ such that

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{2.14}$$

where $\alpha, \beta \in \mathbb{C}$ with $|\alpha|^2 + |\beta|^2 = 1$.

This means that unlike classical bits, for qubits there exist many possible quantum states other than $|0\rangle$ and $|1\rangle$. Moreover, if the state of the qubit is known to be $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, the probability of measuring “0” is $|\alpha|^2$ and the probability of measuring “1” is $|\beta|^2$. This superposition property is one of the primary features of quantum mechanics that distinguishes it from other classical theories.

The vector representations of $|0\rangle$, $|1\rangle$ and $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ are as follows:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad |\psi\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}. \tag{2.15}$$

In Dirac notation the conjugate transpose of each ket is called ‘bra’, $\langle\psi| = (|\psi\rangle)^\dagger$. With this notation, the inner and outer product of two general arbitrary states, $|\psi\rangle = \sum a_j|j\rangle$ and $|\phi\rangle = \sum b_j|j\rangle$(in some basis $\{|j\rangle\}_{j=1}^d$) are defined as follows:

The inner product:

$$\langle\psi|\phi\rangle = \sum_j a_j^* b_j, \tag{2.16}$$

The outer product:

$$|\psi\rangle\langle\phi| = \begin{bmatrix} a_1 \\ \vdots \\ a_d \end{bmatrix} [b_1^* \cdots b_d^*] \,. \tag{2.17}$$

Also, the norm of a ket is defined as

$$|||\psi\rangle|| = \sqrt{\langle\psi|\psi\rangle}. \tag{2.18}$$

Note that all quantum states must have norm one, i.e. $\langle\psi|\psi\rangle = 1$. Therefore, any representation of a general quantum state $|\psi\rangle = \sum a_j|j\rangle$ given in an arbitrary orthonormal basis $\{|j\rangle\}$ satisfies

$$\sum_j |a_j|^2 = 1. \tag{2.19}$$

2.2.1 Density Operator

Typically, we don't have full certainty about the quantum state of a system, meaning that the quantum system is described by an ensemble of quantum states. For instance, the system might be in one of the m states of the set $\{|\psi_j\rangle\}_{j=1}^m$ with some probability p_j. The quantum state of such ensemble is described by its corresponding density operator:

$$\rho = \sum_{j=1}^m p_j|\psi_j\rangle\langle\psi_j|. \tag{2.20}$$

For any density operator $\rho = \sum_{j=1}^m p_j|\psi_j\rangle\langle\psi_j|$, I denote the vector of non-zero eigenvalues by $\boldsymbol{\lambda}(\rho)$, and the set of its corresponding eigenstates by $\{|x\rangle\}$. Furthermore, I assume $\boldsymbol{\lambda}(\rho)$ is sorted in a non-increasing order.

properties of density operators:

1. A density operator ρ is a trace-one, positive semi-definite (Hermitian)

operator that belongs to the space of all linear operators acting on the Hilbert space $\mathcal{H}$, i.e. $\rho \in \mathsf{L}(\mathcal{H})$.

2. Since any ρ is a trace-one Hermitian operator, $\lambda_x(\rho) \geq 0$ for all x, and $\sum \lambda_x(\rho) = 1$.

Note that for a given density operator $\rho = \sum p_j \left|\psi_j\right\rangle \left\langle\psi_j\right|$, the set $\{|\psi_j\rangle\}$ is not unique, and it is not necessarily an orthonormal basis. However, the set $\{|x\rangle\}$ of all eigenstates of a density operator, forms an orthonormal basis, i.e. $\langle x|x'\rangle = \delta(x, x')$.

Example 2.2. The *maximally mixed state* is

$$\rho \equiv \frac{1}{2}\left|0\right\rangle \left\langle 0\right| + \frac{1}{2}\left|1\right\rangle \left\langle 1\right| = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{2.21}$$

It is easy to show that this state can be re-expressed in any arbitrary orthonormal basis $\{|v_0\rangle, |v_1\rangle\}$ by

$$\rho \equiv \frac{1}{2}\left|0\right\rangle \left\langle 0\right| + \frac{1}{2}\left|1\right\rangle \left\langle 1\right| = \frac{1}{2}\left|v_0\right\rangle \left\langle v_0\right| + \frac{1}{2}\left|v_1\right\rangle \left\langle v_1\right|. \tag{2.22}$$

If the actual state of a system is known with certainty, then its density operator has the form $\rho = |\psi\rangle\langle\psi|$ and the state is referred to as a **pure state** [19]. For example, if the system is in the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, its density matrix is:

$$\rho = \begin{bmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{bmatrix}. \tag{2.23}$$

2.2.2 Quantum Measurements

Projective Measurement: One of the primary types of quantum measurements, which has a key role in many QIP applications, is projective measurement. A projective measurement with respect to an arbitrary orthonormal basis $\{|a_x\rangle\}_{x\in\mathcal{X}}$ is defined as follows:

Definition 2.4. A *projective measurement* is represented by a Hermitian operator, called observable $\hat{\mathrm{A}}$, that has a spectral decomposition of the following form:

$$\hat{\mathrm{A}} = \sum_x a_x \Pi_{a_x}, \tag{2.24}$$

where $\Pi_{a_x} = |a_x\rangle \langle a_x|$ is the projector onto the eigenspace of $\hat{\mathrm{A}}$ with eigenvalue a_x.

The eigenvalues a_x represent different outcomes of the measurement. Also, the probability of obtaining x as a result of measuring $|\psi\rangle$ by $\hat{\mathrm{A}}$ is

$$p_x = \langle\psi|\Pi_{a_x}|\psi\rangle. \tag{2.25}$$

Example 2.3. Consider the quantum measurement observable $\hat{\mathrm{Z}} = |0\rangle \langle 0| - |1\rangle \langle 1|$, which measures whether the state is $|0\rangle$ or $|1\rangle$. If the answer is $|0\rangle$ the measurement output is $z_0 = 0$, and if the answer is $|1\rangle$ the measurement output is $z_1 = -1$. Given $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ as the input quantum state to this measurement, the probabilities of obtaining 0 and 1 are:

$$p_0 = \langle\psi|\Pi_0|\psi\rangle = |\alpha|^2,$$
$$p_1 = \langle\psi|\Pi_1|\psi\rangle = |\beta|^2,$$

with $\Pi_0 = |0\rangle \langle 0|$, $\Pi_1 = |1\rangle \langle 1|$.

According to postulates of quantum mechanics, the post-measurement state $|\psi'_x\rangle$ (the state of the system after obtaining m) is the following:

$$|\psi'_x\rangle = \frac{\Pi_{a_x}|\psi\rangle}{\sqrt{p_x}}. \tag{2.26}$$

POVM Measurements: For many applications, we don't care about the post-measurement state of the system, yet we are concerned with calculating probabilities of obtaining different outcomes. POVM measurements are useful for such cases.

Definition 2.5. A *positive operator-valued measurement* (POVM) $\boldsymbol{\mathcal{A}}$ is described by a set $\{\Lambda_m\}_m$ of operators acting on $\mathcal{H}$ satisfying the following:

$$\Lambda_m \geq 0, \quad \forall m, \tag{2.27}$$

$$\sum_m \Lambda_m = \mathbb{1}. \tag{2.28}$$

The probability of obtaining m can be calculated by

$$p_m = \mathrm{Tr}\left\{\Lambda_m \rho\right\}, \tag{2.29}$$

which is equal to $p_m = \langle\psi|\Lambda_m|\psi\rangle$ for a pure state $\rho = |\psi\rangle\langle\psi|$.

2.2.3 Heisenberg Uncertainty Relation

As previously discussed, the superposition principle was one of the first bizarre features of quantum physics. Physicists realized that the outcomes of the measurements on a quantum system do not appear to be predictable; the best they can do is calculating the probabilities of different possible outcomes. This sort of unpredictability could be eliminated by preparing the system in a suitable initial state. In his seminal work [1], Heisenberg introduced a stronger sort of unpredictability: The outcomes of certain pairs of measurements can *never* be predicted simultaneously with certainty—regardless of how the system is prepared! He showed that, for any state preparation, the product of the uncertainties of the position and the momentum of a single quantum particle is bounded from below with a non-zero constant:

$$\Delta x \Delta p \geq \frac{\hbar}{2}. \tag{2.30}$$

Robertson [22] later generalized Heisenberg's formalism to any pair of observables. For general observables $\hat{\mathrm{A}}$ and $\hat{\mathrm{B}}$ the Robertson-Heisenberg uncertainty relation is

$$\Delta\hat{\mathrm{A}}\Delta\hat{\mathrm{B}} \geq \frac{1}{2}|\langle\psi|[\hat{\mathrm{A}}, \hat{\mathrm{B}}]|\psi\rangle|, \tag{2.31}$$

where $\Delta\hat{\mathrm{O}} \equiv \sqrt{\langle\psi|\left(\hat{\mathrm{O}} - \langle\psi|\hat{\mathrm{O}}|\psi\rangle\right)^2|\psi\rangle}$ is the standard deviation of observable $\hat{\mathrm{O}}$, and $[\hat{\mathrm{A}}, \hat{\mathrm{B}}] \equiv \hat{\mathrm{A}}\hat{\mathrm{B}} - \hat{\mathrm{B}}\hat{\mathrm{A}}$ is the commutator of $\hat{\mathrm{A}}$ and $\hat{\mathrm{B}}$.

There are some conceptual shortcomings in Robertson-Heisenberg formulation which motivated development of entropic uncertainty relations (EUR). First, the lower-bound of (2.31) can be zero even if $[\hat{\mathrm{A}}, \hat{\mathrm{B}}] \neq 0$. Second, the left hand side of (2.31) involves variance which can change by mere relabeling [15].

A brief review on EURs, and similar other uncertainty relations will be given in the subsequent chapters.

Chapter 3

Quantifying Uncertainty and Joint Uncertainty

In this chapter, we introduce quantifier-independent notions of uncertainty and joint uncertainty. We use operational information-theoretic principles to identify the common properties of all uncertainty quantifiers. We observe that most existing quantum entropic uncertainty relations use one particular quantifier of joint uncertainty to quantify the uncertainty associated with the two measurements. Our notion relaxes this restriction to one particular quantifier, thereby bringing out previously unexplored joint uncertainty quantifiers. To exemplify the utility of our formalism, we derive a quantum uncertainty relation (in the following chapter) based on one such new quantifier of joint uncertainty.

3.1 Uncertainty and Majorization

Suppose Alice is performing a die role experiment, which we model with a random variable X. We are interested in the uncertainty of X, interpreted as *the minimum uncertainty that Alice necessarily has* prior to the experiment, no matter how much prior knowledge she has. Before doing the experiment, Alice can at most know the distribution $\boldsymbol{p}_X$. Therefore, the uncertainty of X is a property of $\boldsymbol{p}_X$; meaning that, a quantifier of uncertainty is a real-valued function $\mathcal{U}(\boldsymbol{p}_X)$ of probability distributions on X.

Alice can also relabel the faces of her die, for example by replacing each x by $(7-x)$, or even to a different set of labels, such as $\{a, b, \cdots\}$. Let us assume Alice

applies only one-to-one correspondence, i.e. bijective maps. Such transformations are merely a change in Alice's language, and should not make the relabeled outcome Y any more or less uncertain than X. Even if the new alphabet is different from the old one, the effect of any such relabeling on $\boldsymbol{p}_X$ is as though the original labels were just permuted amongst themselves:

$$\boldsymbol{p}_Y = P_\pi \boldsymbol{p}_X,$$

where P_π is the matrix representation of a permutation π of the elements of X. Thus, from the uncertainty standpoint, permutations of a fixed alphabet of labels effectively capture all relabeling schemes. Thus far, we have found a way to tell when the uncertainties of two variables are equal. That is if $\boldsymbol{p}_Y = P_\pi \boldsymbol{p}_X$, then $\mathcal{U}(\boldsymbol{p}_Y) = \mathcal{U}(\boldsymbol{p}_X)$.

Now imagine the following thought experiment. Alice tosses a fair coin before rolling her die. If she gets tails she will relabel the faces of her die with a permutation π_1 and if she gets heads she will relabel the faces of her die with a permutation π_2. Then, right before rolling the die, she forgets the coin toss result and thus does not know which permutation she has applied. In this case, from Alice's perspective, the probability distribution of the die roll experiment after forgetting the coin toss result, is

$$\boldsymbol{p}_Y = \frac{1}{2} P_{\pi_1} \boldsymbol{p}_X + \frac{1}{2} P_{\pi_2} \boldsymbol{p}_X, \tag{3.1}$$

which describes a more uncertain random variable.

In general, we can conclude that the resulting random variable after any random relabeling followed by forgetting is more uncertain than the initial random variable. In other words, any uncertainty quantifiers must be monotonically non-decreasing under the action of all random relabelings followed by forgetting.

Note that the action of a random relabeling followed by forgetting on a probability

distribution is equivalent to the action of a convex combination of permutations. As discussed in chapter 2, any convex combination of permutation matrices is equal to a doubly stochastic matrix. Recall that if $\boldsymbol{p}_Y = D\boldsymbol{p}_X$, where D is doubly stochastic matrix, then $\boldsymbol{p}_Y \prec \boldsymbol{p}_X$. Thus all Schur-concave functions (functions that preserve majorization relations) are qualified to be valid quantifiers of uncertainty [15].

In conclusion, an experiment modeled with probability distribution $\boldsymbol{p}_Y$ is more uncertain than an experiment with associated probability distribution $\boldsymbol{p}_X$ if and only if $\boldsymbol{p}_Y \prec \boldsymbol{p}_X$.

3.2 Not All Doubly Stochastic Maps Are Allowed !

As we saw in the previous section, for experiments such as a die roll, the uncertainty quantifier has to be monotonic under the action of any doubly stochastic map. In general, we show that not all doubly stochastic maps are allowed.

We consider two different information-theoretic thought experiments and show that a valid quantifier of uncertainty has to be monotonic under the action of a subset of all doubly stochastic matrices. The first scenario is similar to the random relabeling experiment discussed before.

3.2.1 Restrictions Due to Physical Symmetries

In Alice's random relabeling and forgetting experiment discussed earlier, any permutation π results in a variable Y that, despite being different from X, still carries the meaning of a die-roll outcome. Instead, if X were a physical observable such as the energy of a quantum harmonic oscillator, arbitrary permutations could result in new variables that can no longer be interpreted as energy levels. In this sense, to preserve the physical meaning of energy represented by the random variables, we would have

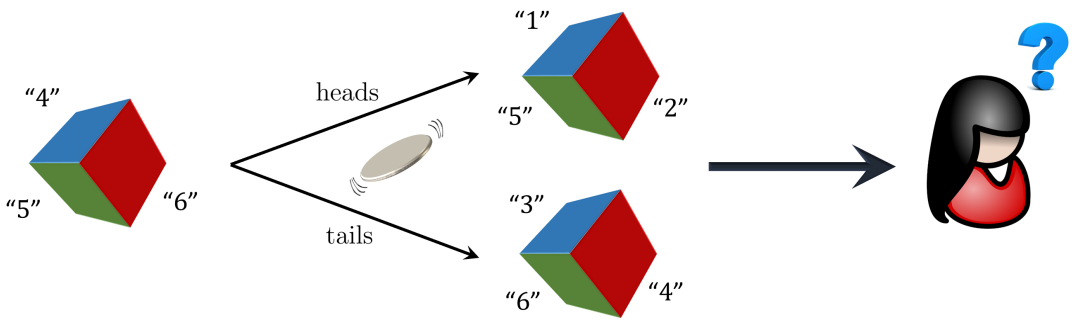


Figure 3.1: Alice has a die and knows its probability distribution. She randomly relabels her die based on a coin toss. Then she forgets the coin toss result, which makes her more uncertain about the outcome of her die.

to restrict the allowed permutations to only shifts. In general, the restricted class of relabelings is a group G of *symmetries* of the observable underlying the random variable, with each symmetry $g \in G$ corresponding to a *change in one's reference frame*. For discrete observables, G is a subgroup of the group of all permutations, while for observables with uncountably many possible values (e.g. positions and momenta of particles) the symmetries g are transformations of continuous variables (e.g. Galilean transformations).

Our first requirement from a valid quantifier $\mathcal{U}$ of uncertainty is that it must be invariant under the symmetry group G of the underlying observable:

$$\mathcal{U}(P_g \boldsymbol{p}_X) = \mathcal{U}(\boldsymbol{p}_X). \tag{3.2}$$

This immediately leads to the following: two variables X and Y, both representing the same physical observable, are considered *equally uncertain* if their distribution vectors are related by some permutation $g \in G$, i.e. $\boldsymbol{p}_Y = P_g \boldsymbol{p}_X$.

Now consider Alice's random relabeling and forgetting experiment discussed previously (see Fig. 3.1). Recall that Alice tosses a coin before rolling her die; she then relabels the die's faces with a permutation that is determined by the outcome of this coin toss; then she forgets the result of the coin toss. The outcome Y of the relabeled

die in this modified experiment cannot be more certain than X—after all, X carries the fundamental uncertainty of the die. In fact, the random choice of relabeling (about which Alice is a priori ignorant) renders Y *more* uncertain than X.

Suppose $g \in G$ is a permutation chosen at random from the symmetry group G under some distribution $\boldsymbol{t} \equiv (t_g)$. If variable X is transformed under this random action of g, then the resulting random variable Y is distributed as

$$\boldsymbol{q}_Y = D^{\text{sym}} \boldsymbol{p}_X, \tag{3.3}$$

where $D^{\text{sym}} = \sum_{g \in G} t_g P_g$ is the convex combination of the symmetry matrices P_g under the distribution $\boldsymbol{t}$. Since each P_g is a permutation matrix, every possible D^{sym} is doubly stochastic. We denote by $\mathcal{D}^{\text{sym}}$ the set of all such D^{sym} matrices.

3.2.2 Restrictions Due to Information Processing

Now, consider a different scenario: Alice performs an experiment whose random output is modeled by random variable X. She sends the outcome X to Bob via some classical channel given by the column-stochastic matrix $T \equiv (T_{y|x})$. Bob will get information about X in the form of a new random variable Y, which in turn is distributed according to $\boldsymbol{q}_Y = T\boldsymbol{p}_X$. Using his knowledge of the channel, Bob plans to recover X (see Fig. 3.2).

Why does Bob need to recover X from the channel output? Because the uncertainty of Y, the channel output, will not necessarily have a meaningful relation to the uncertainty of X. The channel could transmit x perfectly, or with some added noise. In these cases, the output Y is equally or more uncertain than X. On the other hand, the channel could also completely ignore x and output some constant value, in which case the uncertainty of Y would be *less* than that of X. Actually, the channel

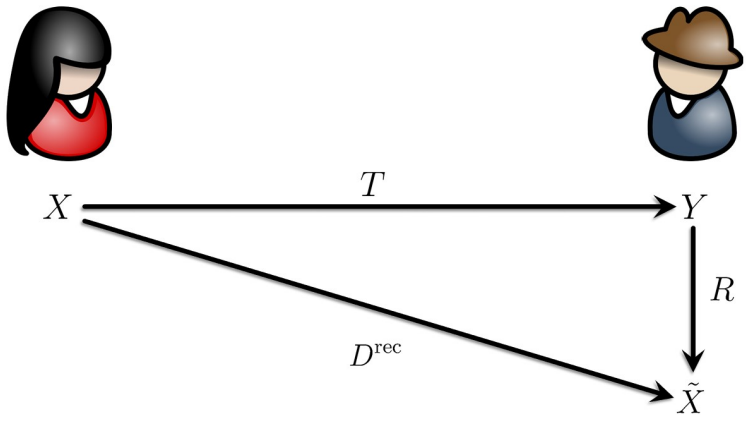


Figure 3.2: Bob tries to recover Alice's message about the variable X after it has been corrupted by the channel T.

might result in information in a fundamentally different form from X. For example, the channel could just send the parity of the die outcome to Bob, in which case Y does not even represent the same underlying observable as X. Hence, we cannot make a general statement about the relation between the uncertainty of X and the uncertainty of Y. In order to compare the uncertainty of X with the uncertainty of a relevant random variable, we must extract from Y some variable that has the same physical meaning as X, so that we can treat them both on an equal footing.

In this scenario, Bob's aim in his recovery task is not to maximize his chances of guessing X correctly, but rather to faithfully account for the uncertainty that Y contains about X. Given that Bob does not have any access to $\boldsymbol{p}_X$, one natural recovery protocol would be to do a randomized likelihood guessing. In this recovery method, Bob's guess for the likelihoods of different x's that could have resulted in a particular final label y are of the form $c_y T_{y|x}$, where c_y is a proportionality constant depending only on y. Therefore, the corresponding recovery scheme is of the following

form

$$R_{x|y} = \frac{T_{y|x}}{\sum_{x'} T_{y|x'}}. \tag{3.4}$$

The resulting distribution of $\tilde{X}$ is given by the composite action of T and R on $\boldsymbol{p}_X$:

$$\boldsymbol{p}_{\tilde{X}}^{\text{rec}} = RT\boldsymbol{p}_X =: D^{\text{rec}}\boldsymbol{p}_X. \tag{3.5}$$

For any column-stochastic T, with the corresponding R (note that R is completely determined by T) constructed as in (3.4), the matrix $D^{\text{rec}} = RT$ is doubly stochastic. In fact, $D^{\text{rec}} = (d_{jk})$ where $d_{jk} = \sum_y R_{j|y}T_{y|k}$. Thus

$$\sum_j d_{jk} = \sum_j \sum_y \frac{T_{y|j}T_{y|k}}{\sum_x T_{y|x}} \tag{3.6}$$

$$= \sum_y \frac{\sum_j T_{y|j}T_{y|k}}{\sum_x T_{y|x}} \tag{3.7}$$

$$= \sum_y T_{y|k} = 1, \tag{3.8}$$

and analogously $\sum_k d_{jk} = 1$, which proves doubly stochasticity of D^{rec}.

The (necessarily degenerative) evolution of the uncertainty of a particular variable, under *any* classical processing (represented by the action of the channel) of the variable, is always via such matrices. We denote the collection of all such matrices by $\mathcal{D}^{\text{rec}}$.

While $\mathcal{D}^{\text{sym}}$ depends on the symmetry group of the observable, $\mathcal{D}^{\text{rec}}$ depends only on the size of the outcome space. For a variable with complete permutation symmetry, as noted above, $\mathcal{D}^{\text{sym}}$ contains all doubly stochastic matrices, in particular all of $\mathcal{D}^{\text{rec}}$. But in situations with restricted symmetries, each class can contain members not belonging to the other. For instance, consider a variable with 3 possible outcomes and whose symmetry group is the (order-3) group of cyclic permutations. The two nontrivial permutations are transformations contained (by design) in $\mathcal{D}^{\text{sym}}$, but not

in $\mathcal{D}^{\text{rec}}$. On the other hand, we can show that the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \end{pmatrix}$$

is in $\mathcal{D}^{\text{rec}}$, but not in $\mathcal{D}^{\text{sym}}$. Therefore, the structure of the union of these classes cannot be reduced to either one of the classes. This example can be generalized naturally to higher dimensions. Further characterization of the classes $\mathcal{D}^{\text{rec}}$ and $\mathcal{D}^{\text{sym}}$ is an interesting problem that we leave for future work.

Recall that majorization relations are induced by the action of all doubly stochastic matrices. Now the interesting question is to find and characterize the hierarchy relations among real vectors induced by the action of doubly stochastic matrices in the union of $\mathcal{D}^{\text{rec}}$ and $\mathcal{D}^{\text{sym}}$. We also leave this question here for future researches.

3.3 Single-Variable Quantifiers of Uncertainty

The "sym" and "rec" classes of doubly stochastic matrices together capture the mechanisms of uncertainty increase in our formalism. Based on this, we have the following definition:

Definition 3.1. Let random variable X with alphabet $\mathcal{X}$ and dimension $d = |\mathcal{X}|$ represent a physical system with underlying symmetry group G. Also let P_g denote the matrix representation of permutation $g \in G$, and $\mathcal{D}^{\text{sym}} = \{D | D = \sum_{g \in G} t_g P_g, \quad \forall \boldsymbol{t} = (t_g)\}$ and $\mathcal{D}^{\text{rec}} = \{D | D = RT, \quad \forall T\}$, where $\boldsymbol{t}$ is a probability vector, and T is a matrix representation of a classical channel with d inputs and some arbitrary number of outputs, and R is the matrix representation of the recovery map based on T defined as in (3.4). Then, a *quantifier of uncertainty* of the random variable X is a function

$\mathcal{U}$ of the distribution $\boldsymbol{p} \equiv \boldsymbol{p}_X$ of the variable, satisfying the following constraints:

(i) Invariance under symmetry group G

$$\mathcal{U}(P_g \boldsymbol{p}) = \mathcal{U}(\boldsymbol{p}), \quad \forall g \in G. \tag{3.9}$$

(ii) Monotonicity under random symmetry-transformations:

$$\mathcal{U}(D\boldsymbol{p}) \geq \mathcal{U}(\boldsymbol{p}), \quad \forall D \in \mathcal{D}^{\text{sym}}. \tag{3.10}$$

(iii) Monotonicity under information processing and recovering:

$$\mathcal{U}(D\boldsymbol{p}) \geq \mathcal{U}(\boldsymbol{p}), \quad \forall D \in \mathcal{D}^{\text{rec}}. \tag{3.11}$$

Here "*sym*" is determined by the symmetries of the variable's underlying physical observable. Without loss of generality, we will also require $\mathcal{U}(\boldsymbol{e}) = 0$, where $\boldsymbol{e} \equiv (1, 0, 0 \ldots, 0)$.

A function satisfying (i) and (ii), but not (iii), will be called a *weak quantifier of uncertainty*, since the "sym" class is more important than the "rec". Indeed, the former is based on the natural symmetries of an observable, and therefore the constraints that it induces on uncertainty quantifiers are inviolable. On the other hand, "rec", even though it is an essential ingredient in the strictest information-theoretic definition of uncertainty, can be ignored in natural situations where information processing is not involved. Functions that respect the "sym" constraints, but violate the "rec" ones nevertheless turn out to be useful indicators of uncertainty.

Claim 3.1. *All Schur-concave functions are valid uncertainty quantifiers under definition 3.1.*

If the observable's symmetry group G includes all permutations, then by Birkhoff's theorem [21, 23] $\mathcal{D}^{\text{sym}}$ is the set of *all* doubly stochastic matrices. Hence, condition

(3.10) implies (3.11), because D in (3.10) can be *any* doubly stochastic matrix. In this case, any Schur-concave function, e.g. Shannon entropy and Rényi entropies, can play the role of an uncertainty quantifier [17]. In other words, the majorization relation $\boldsymbol{q}_Y \prec \boldsymbol{p}_X$ indicates that the uncertainty of Y is more than X. Note that, as the majorization ordering is not total, no single Schur-concave function can determine the relation between two arbitrary vectors under this ordering [15].

If a variable has restricted symmetries, then the uncertainty hierarchy of its distributions becomes different from the majorization hierarchy. All Schur-concave functions still remain valid uncertainty quantifiers. But in addition, by virtue of the reduction in the class $\mathcal{D}^{\text{sym}}$, some non-Schur-concave functions could also qualify to be quantifiers of uncertainty. Not all such quantifiers may respect condition (iii), which makes them *weak* quantifiers of uncertainty.

Claim 3.2. *The variance and standard deviation are weak quantifiers of uncertainty.*

In fact, we prove the following: If the observable underlying X can take integer values, and if its symmetries are shift permutations, the variance of X is only a weak uncertainty quantifier (i.e. does not satisfy (iii)). An example of such an observable is the energy of a quantum Harmonic oscillator.

Let X be a variable that takes positive integer values, $\{x \in \mathbb{Z}_+\}$ (any variable with equally-spaced possible values can be suitably rescaled). As in the case of the energy levels of a harmonic oscillator, we will consider the symmetries of X to be shifts, i.e. transformations $x \mapsto x + d$, for integers d. Let's investigate if the variance σ_X^2 is a valid quantifier of the uncertainty of X. To this end, we must check if the variance is monotonically non-decreasing under the two classes of doubly stochastic matrices, $\mathcal{D}^{\text{rec}}$ and $\mathcal{D}^{\text{sym}}$. First, unsurprisingly, we show that the variance is non-decreasing under $\mathcal{D}^{\text{sym}}$, i.e. the class of all random shifts.

Let a shift by d_k be applied with probability q_k. Then X is replaced by a new random variable Y with realizations $x - d_k$, and associated probability distribution $\boldsymbol{p}_Y = \boldsymbol{p}_X \otimes \boldsymbol{q}_K$, which has the following variance:

$$\begin{aligned}
\sigma_Y^2 =& \mathbb{E}\{Y^2\} - \mathbb{E}\{Y\}^2 \\
=& \sum_{x,k} p_x q_k (x - d_k)^2 - \left(\sum_{x,k} p_x q_k (x - d_k) \right)^2 \\
=& \mathbb{E}\{X^2\} + \mathbb{E}\{K^2\} - 2\mathbb{E}\{X\}\mathbb{E}\{K\} - (\mathbb{E}\{X\} - \mathbb{E}\{K\})^2 \\
=& \mathbb{E}\{X^2\} - \mathbb{E}\{X\}^2 + \mathbb{E}\{K^2\} - \mathbb{E}\{K\}^2 \\
=& \sigma_X^2 + \sigma_K^2.
\end{aligned}$$

Therefore, $\sigma_Y^2 \geq \sigma_X^2$, as we had anticipated. Now we provide a counterexample that shows that the variance can decrease under $\mathcal{D}^{\text{rec}}$. Consider a channel T that is the identity channel except for the block acting on $x \in \{0, 1\}$, where it acts as the completely-randomizing channel:

$$T = \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & \\ 0 & & \mathbb{1} \end{pmatrix}. \tag{3.12}$$

By a straightforward calculation one can show that the recovery channel would also have the same matrix representation, $R = T$, resulting in the overall transformation $D^{\text{rec}} = R = T$ on the probability distribution. Now consider an initial distribution $\boldsymbol{p}_X$ such that $p_0 = p_2 = 0.5$, with all other components zero; thus the initial variance is $\sigma_X^2 = 1$. The outcome of the recovery channel would be a random variable $\tilde{X}$ with probability vector $\boldsymbol{p}_{\tilde{X}}^{\text{rec}}$ such that $p_2^{\text{rec}} = 0.5$ and $p_0^{\text{rec}} = p_1^{\text{rec}} = 0.25$. As we can see, $\sigma_{\tilde{X}}^2 = 11/16 < \sigma_X^2 = 1$.

This shows that the variance can be decreased under $\mathcal{D}^{\text{rec}}$, and is therefore only a weak uncertainty quantifier under our definition.

3.4 Joint Uncertainty

The uncertainty of the outcomes of *individual* measurements cannot provide a complete description of the uncertainty principle since uncertainty relations are characterized in terms of lower bounds on quantifiers of *joint* uncertainty. The joint uncertainty is about the outcomes of at least two measurements. For example, in Heisenberg's uncertainty relation the joint uncertainty of the position and momentum of a particle is measured through the quantity $\Delta x \Delta p$. Other uncertainty relations use various other quantifiers of joint uncertainty, e.g. the sum of Rényi entropies of the outcome distributions of individual experiments [4]. We now formulate a notion that captures the *essence* of the concept of joint uncertainty.

In general, we seek a notion of the joint uncertainty of a collection of n potential experiments, each of which could have an arbitrary number of possible outcomes. For simplicity, here we will restrict the discussion to collections of $n = 2$ experiments, each with a finite number of possible outcomes. Generalization to arbitrary n can be carried out in a straightforward manner. Before we state the precise definition of joint uncertainty, we give two examples to demonstrate the different forms that joint uncertainty can take in different scenarios.

Example 3.1. Consider two experiments, one tossing a coin whose 2 outcomes are represented by the random variable X, and the other is rolling a die whose 6 outcomes are represented by the random variable Y. To define the joint uncertainty of the two experiments, we construct a new combined experiment as an independent and simultaneous performance of *both* the original experiments. In other words, tossing the coin and independently rolling the die (see Fig. 3.3). In this scenario, the joint uncertainty of the outcomes X and Y, is the *single-variable* uncertainty of the combined

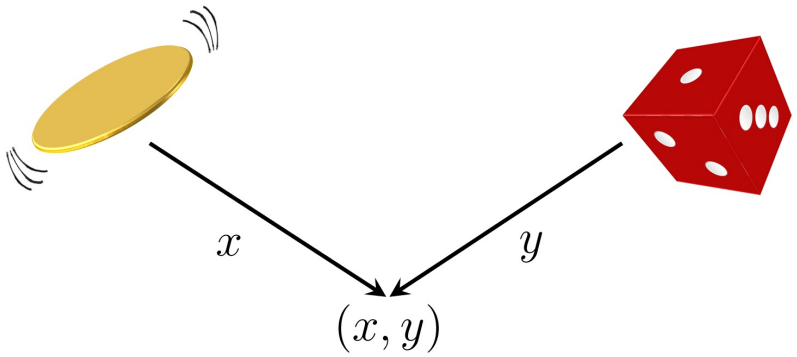


Figure 3.3: Example 1: Both experiments are performed simultaneously. The probability to obtain the combined outcome $z \equiv (x, y)$ is $p_{Z=z} = p_{X=x} p_{Y=y}$.

variable $Z \equiv (X, Y)$. Note that the combined variable Z has $|\mathcal{Z}| = |\mathcal{X}||\mathcal{Y}| = 12$ outcomes. In this case, the probability distribution of the combined variable Z is given by $\boldsymbol{p}_Z = \boldsymbol{p}_X \otimes \boldsymbol{p}_Y$, and therefore the joint uncertainty can be written as

$$\mathcal{U}(\boldsymbol{p}_X \otimes \boldsymbol{p}_Y), \tag{3.13}$$

where $\mathcal{U}$ is a quantifier of single-variable uncertainty as per Def. 3.1. Most joint uncertainty quantifiers considered in the literature, e.g. the sum of Shannon entropies of the individual outcome distributions, can be interpreted as a quantifier of uncertainty of such a combined experiment [15].

Example 3.2. In this example we consider again the same two random variables X and Y. But this time we first toss a different coin to make a choice between the actions "toss the coin" (resulting in outcome X) and "roll the die" (leading to Y), and then perform only the chosen action (Fig. 3.4). The combined variable Z of this experiment has $|\mathcal{Z}| = |\mathcal{X}| + |\mathcal{Y}| = 8$ possible outcomes, and the uncertainty of the outcome of this new experiment is (modulo the uncertainty in the choice of action) also a manifestation of the joint uncertainty of (X, Y). In this case, if the first coin

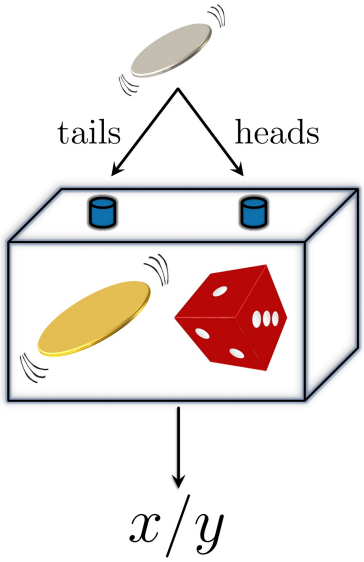


Figure 3.4: Example 2: Another combination of two experiments; only one of them is performed.

is unbiased, $\boldsymbol{p}_Z = \frac{1}{2}\boldsymbol{p}_X \oplus \frac{1}{2}\boldsymbol{p}_Y$ and therefore the joint uncertainty can be written as

$$\mathcal{U}\left(\frac{1}{2}\boldsymbol{p}_X \oplus \frac{1}{2}\boldsymbol{p}_Y\right)$$

where $\mathcal{U}$ is a single-variable uncertainty quantifier. The quantifiers of joint uncertainty proposed in [7] can be interpreted as quantifiers of uncertainty of such a single combined experiment.

As these scenarios illustrate, there could be different ways in which experiments could be combined into one super-experiment, whose outcomes have uncertainty that reflects an aspect of the joint uncertainty of (X, Y). But the essence of joint uncertainty is not quite captured by any one of these joint experiments. In fact, the functions

$$H_\alpha(\boldsymbol{p}_X) + H_\beta(\boldsymbol{p}_Y), \tag{3.14}$$

which are quantifiers of joint uncertainty proposed in [4] (where H_α and H_β are Rényi entropies), cannot be interpreted as a quantifier of the uncertainty of *any* single

combined experiment (unless $\alpha = \beta$). Another such quantifier is the Heisenberg quantity $\Delta x \Delta p$.

Our ultimate intention is to apply the notion of joint uncertainty to the quantum uncertainty principle, which is simply about several *potential measurements*, each a potential (actual or counterfactual) experiment in its own right. By all means, we could construct counterfactual scenarios that feature two or more of these measurements, but the uncertainty principle itself neither depends on, nor is bound by, such "connected narratives".

These considerations indicate that the notion of joint uncertainty is not bound to a single combined experiment, or even to the concept of combined experiments. Let us identify the desired properties of a quantifier of the joint uncertainty of two random variables X and Y. The pairs (X, Y) that have the smallest joint uncertainty are ones where both distributions are completely certain. The most jointly-uncertain pairs, on the other hand, are the ones where both variables are completely uncertain. Furthermore, all the quantifiers of the joint uncertainty of (X, Y) are real-valued functions of the distributions $\boldsymbol{p} \equiv \boldsymbol{p}_X$ and $\boldsymbol{q} \equiv \boldsymbol{q}_Y$, and must reduce to the quantifiers of single-variable uncertainty (as in Def. 3.1) if one of the vectors $\boldsymbol{p}$ and $\boldsymbol{q}$ is kept fixed. This brings us to the following definition:

Definition 3.2. A quantifier of joint uncertainty of two variables X and Y is a real-valued function $\mathcal{J}$ of $(\boldsymbol{p}, \boldsymbol{q}) \equiv (\boldsymbol{p}_X, \boldsymbol{q}_Y)$, such that

$$\mathcal{J}\left(D_1\boldsymbol{p}, D_2\boldsymbol{q}\right) \geq \mathcal{J}\left(\boldsymbol{p}, \boldsymbol{q}\right) \tag{3.15}$$

for all doubly stochastic matrices D_1, D_2 in the respective "*sym*" and "*rec*" classes of both variables. Without loss of generality, we can require that $\mathcal{J}\left(\boldsymbol{e}^1, \boldsymbol{e}^2\right) = 0$ (where $\boldsymbol{e}^1$ and $\boldsymbol{e}^2$ are maximally-certain distributions on the respective spaces), so that

$\mathcal{J}(\boldsymbol{p}, \boldsymbol{q})$ is always nonnegative. As in the single-variable case, we will call functions satisfying (3.15) for the "*sym*" class, but not for the "*rec*" class, *weak quantifiers of joint uncertainty.*

Claim 3.3. *All entropic quantifiers of uncertainty including the one in* (3.14) *are all valid joint uncertainty quantifiers under definition 3.2. The* $\Delta x \Delta p$ *quantifier in Heisenberg's relation is a weak joint uncertainty quantifier under a restriction to (random) Galilean transformations.*

Note that if the symmetry groups of both variables are the respective full permutation groups, then D_1 and D_2 can be *any* two doubly stochastic matrices of appropriate dimensions. In this case, the relation in (3.15) states that $\mathcal{J}$ is monotonic under the *direct product relation* "$\gg$" defined by:

$$(\boldsymbol{p}_1, \boldsymbol{q}_1) \gg (\boldsymbol{p}_2, \boldsymbol{q}_2) \Leftrightarrow (\boldsymbol{p}_1 \succ \boldsymbol{p}_2 \text{ and } \boldsymbol{q}_1 \succ \boldsymbol{q}_2).$$

An important advantage of Definition 3.2 is that it gives us the chance to find more, new quantifiers of joint uncertainty. As we discussed, quantifiers of joint uncertainty should neither be necessarily interpreted with a practical setup of actual experiments nor be a mathematical combination of different single-variable uncertainty quantifiers, such as product of two quantifiers as in $\Delta x \Delta p$ or sum of two different entropies as in $H_\alpha(\boldsymbol{p}_X) + H_\beta(\boldsymbol{p}_Y)$. To show the existence of joint uncertainty quantifiers independent of such restrictions, we introduce the following function:

$$\mathcal{J}_2(\boldsymbol{p}, \boldsymbol{q}) = 1 - \boldsymbol{p}^{\downarrow} \cdot \boldsymbol{q}^{\downarrow}, \tag{3.16}$$

where $\boldsymbol{p}^{\downarrow}$ is the rearranged version of $\boldsymbol{p}$ in a non-increasing order, and $(\cdot)$ denotes the usual dot product.

To show that $\mathcal{J}_2(\boldsymbol{p}, \boldsymbol{q})$ is a valid quantifier of joint uncertainty, we need to prove that it is a Schur-concave function of $\boldsymbol{p}$ for a fixed distribution $\boldsymbol{q}$. Assume that $\mathcal{J}_2 = 1 - (\cdots + p_k q_l + \cdots + p_j q_m + \cdots)$, with $p_k \geq p_j$ and thus $q_l \geq q_m$. Then:

$$(p_k - p_j)\left(\frac{\partial \mathcal{J}_2}{\partial p_k} - \frac{\partial \mathcal{J}_2}{\partial p_j}\right) =$$

$$(p_k - p_j)(-q_l + q_m) \leq 0.$$

It is apparent that $\mathcal{J}_2$ is also invariant under different permutations of $\boldsymbol{p}$, as we have to sort $\boldsymbol{p}$ and $\boldsymbol{q}$ initially and then compute $\boldsymbol{p}^{\downarrow} \cdot \boldsymbol{q}^{\downarrow}$. Therefore, according to Theorem 2.4 we have proved that $\mathcal{J}_2$ is a Schur-concave function of $\boldsymbol{p}$ for a fixed distribution $\boldsymbol{q}$. Analogously we can show that $\mathcal{J}_2$ is also a Schur-concave function of $\boldsymbol{q}$ for a fixed distribution $\boldsymbol{p}$ and thus, is indeed a valid quantifier of joint uncertainty under definition 3.2.

3.5 Conclusion

In this chapter, we discovered a unifying axiomatic definition of a valid uncertainty quantifier. More importantly, we showed that variance, which seemed to be fundamentally different from all entropy functions, can be defined as an appropriate (but weak) quantifier of uncertainty, under the very same defining characterizations of entropies as (strong) valid uncertainty quantifiers.

By taking the same approach, we defined valid joint uncertainty quantifiers. The novelty of this definition is its independence from different practical scenarios with which most existing joint uncertainty quantifiers are defined. Though some joint uncertainty quantifiers, such as Heisenberg's $\Delta x \Delta p$, cannot be interpreted with a single combination of different actual measurements, they are mathematically constructed from single-variable uncertainty quantifiers. Our definition of joint uncertainty quan-

tifiers is also free of individual single-variable quantifiers of uncertainty. To show that these usual approaches to define joint uncertainty quantifiers are unnecessarily restrictive, we gave one simple example of a joint uncertainty quantifier, which neither can be interpreted with a practical setup of experiments nor is a mathematical combination of single-variable quantifiers of uncertainty.

Chapter 4

Fine-Grained Quantum Uncertainty Relations

In this chapter, we utilize the notion of joint uncertainty developed in the previous chapter to show how our formalism generalizes a large class of uncertainty relations. There are two major types of preparational uncertainty relations: those based on standard deviation, and entropic uncertainty relations (EURs). A very brief review of the early history of these URs is given in the next section. Thereafter, we show that our formalism not only unifies these two large classes of uncertainty relations, but also provides a method for deriving new uncertainty relations. We take the joint uncertainty quantifier, which was introduced in chapter 3, and prove that these new quantum uncertainty relations exist. This relation is non-trivially different from all existing URs.

We further use our formalism to study the theory of fine-grained quantum uncertainty relations. Fine-grained uncertainty relations (or quantifier-free URs) are mathematical expressions that are formulated independent of any particular uncertainty quantifier. We see that there exist no universal quantifier-free formalism of all uncertainty relations. However, under some restrictions fine-grained uncertainty relations can be found. Using majorization techniques, we then prove a new class of fine-grained uncertainty relations.

4.1 Background

Revealing one of the most striking features of quantum mechanics, Heisenberg [1] showed that the outcomes of certain pairs of measurements on a quantum system can never be predicted simultaneously with certainty—regardless of how the system is prepared. Heisenberg's original statement of what he called the "indeterminacy" principle concerned potential measurements of the position and the momentum of a single quantum particle.

Later, Robertson [22] and others [24–27] generalized Heisenberg's original idea, both in the number and type of measurements involved, and in the quantifiers used to quantify joint uncertainty. At the same time, Heisenberg himself set off another chain of research on a related concept: measurement-induced disturbance and so-called noise-disturbance relations [28–31].

The Heisenberg–Robertson class of uncertainty relations use the standard deviation as a quantifier of the uncertainty of the outcomes of a measurement:

$$\Delta\hat{\mathrm{A}}\Delta\hat{\mathrm{B}} \geq \frac{1}{2}|\langle\psi|[\hat{\mathrm{A}},\hat{\mathrm{B}}]|\psi\rangle|, \tag{4.1}$$

where $\Delta\hat{\mathrm{O}} \equiv \sqrt{\langle\psi|\left(\hat{\mathrm{O}} - \langle\psi|\hat{\mathrm{O}}|\psi\rangle\right)^2|\psi\rangle}$ is the standard deviation of observable $\hat{\mathrm{O}}$. However, one can use other quantifiers of uncertainty to formulate the quantum mechanical indeterminacy principle. Hirschman [25] introduced the first uncertainty relation in terms of *entropies*. Later, improved versions of Hirschman's entropic relation, proved in [26], were shown to imply Heisenberg's original statement [27]. Deutsch [32] provided an objective argument for the superiority of entropic uncertainty relations to the previous variance-based approach. First, he argued that the lower bound of (4.1) depends on the state of the system $|\psi\rangle$. Thus, the lower bound

can be trivially zero if $|\psi\rangle$ belongs to the null space of commutator $[\hat{\mathrm{A}}, \hat{\mathrm{B}}]$, which is possible in finite dimensions. Second, he indicated that the variance quantifiers, such as standard deviation used in Heisenberg's relation, can be increased by mere relabeling of the random variables associated with the measurement results. He concluded that the essence of uncertainty is contained in the probability distributions of the outcomes, rather than the somewhat arbitrary *values* that we associate with them. He also found a state-independent entropic uncertainty relation, which Maassen and Uffink [4] later improved to:

$$H_1(\boldsymbol{p}_X) + H_1(\boldsymbol{q}_Y) \geq -2\log\eta, \tag{4.2}$$

where $\hat{\mathrm{A}} = \sum_x a_x |a_x\rangle \langle a_x|$ and $\hat{\mathrm{B}} = \sum_y b_y |b_y\rangle \langle b_y|$ are two projective measurements with associated probability vectors $\boldsymbol{p}_X$ and $\boldsymbol{q}_Y$, respectively; and η is the maximum overlap of the two measurements defined as $\eta \equiv \max_{x,y} |\langle a_x|b_y\rangle|$.

Deutsch's work pioneered a paradigm shift towards more information-theoretic quantifiers of uncertainty [33–37]. Many different entropic relations have since been proposed, mostly using Shannon and Rényi entropies as quantifiers of uncertainty [38, 39]. Much effort has also been directed towards generalizing the concept to more than two measurements [40–42].

Moreover, a new type of URs have been introduced recently: fine-grained uncertainty relations. Instead of setting the constraints on the output probability distributions of the given set of measurements through some inequalities in terms of uncertainty quantifiers, these quantifier-free URs are mathematically formulated directly in terms of the output probability distributions themselves (see [7, 14–16, 43, 44]).

Despite the increasing research trend towards tightening the bounds of different EURs and finding new applications of them in other research areas, here we focus on a better understanding of the essence of the uncertainty principle itself.

4.2 The Quantum Uncertainty Principle

According to the laws of quantum mechanics, given a pair of incompatible quantum measurements (non-commuting observables), it is impossible to prepare a quantum state such that both measurements result in certain outcomes. In this case, at least one measurement outcome is always uncertain to some degree. Generally, for any set of two or more measurements, quantum mechanics predicts restrictions on the probability distributions associated with the measurements outcomes of on an arbitrary quantum state. The "uncertainty principle" is indeed a collection of identities known as uncertainty relations (URs), all quantifying these restrictions.

Broadly, there are three different operational contexts of URs: different measurements applied on the same quantum state (either counterfactually or by preparing many copies of the same state); (approximately) simultaneous execution of several measurements; and sequential execution of several measurements. The notions that we developed in the last two sections can be applied in all of these contexts, since they all include instances of finite-dimensional classical variables. Here we will focus on the first type of situation, where different measurements are considered on identical preparations.

Since these URs depend only on the probabilities of the outcomes, a positive-operator–valued measure (POVM) description of measurements is adequate in the formalism. Consider the case of two POVM's $\boldsymbol{\mathcal{A}} \equiv \{\Lambda_x\}_x$ and $\boldsymbol{\mathcal{B}} \equiv \{\Gamma_y\}_y$. Note that without loss of generality we assume both $\boldsymbol{p}$ and $\boldsymbol{q}$ have the same length d; we can always pad some zeros to the shorter vector. For a quantum state ρ, measurement $\boldsymbol{\mathcal{A}}$ leads to outcome probability distribution $\boldsymbol{p}(\rho)$ where $p_x(\rho) = \mathrm{Tr}\{\Lambda_x \rho\}$, and $\boldsymbol{\mathcal{B}}$ to $\boldsymbol{q}(\rho)$ with $q_y(\rho) = \mathrm{Tr}\{\Gamma_y \rho\}$.

For a so-called *incompatible* pair of POVM's $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}})$, there is no ρ that results in both $\boldsymbol{p}(\rho)$ and $\boldsymbol{q}(\rho)$ completely certain, leading to the existence of a "minimal joint uncertainty". Many URs are statements to this effect:

$$\mathcal{J}\left(\boldsymbol{p}(\rho), \boldsymbol{q}(\rho)\right) \geq c \quad \forall \rho, \tag{4.3}$$

where $\mathcal{J}$ is a quantifier of joint uncertainty, and $0 < c \leq C_{\mathcal{J}}\left(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}}\right) := \min_{\rho} \mathcal{J}\left(\boldsymbol{p}(\rho), \boldsymbol{q}(\rho)\right)$. In some relations (e.g. Robertson's), c is not a constant but rather a non-negative function of ρ. The disadvantage of such a state-dependent lower bound is that it can be zero in some cases even if $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$ are incompatible. For this reason, state-independent c's are favored in most of the recent literature.

Claim 4.1. *In general, our analysis of uncertainty and joint uncertainty enables us to unify the understanding of all URs of the form*

$$\mathcal{J}\left(\boldsymbol{p}_1(\rho), \boldsymbol{p}_2(\rho) \ldots, \boldsymbol{p}_n(\rho)\right) \geq c, \quad \forall \rho \tag{4.4}$$

where $\mathcal{J}(\cdot)$ is a (strong or weak) joint uncertainty quantifier (under a generalized version of Def. 3.2) of the n probability distributions $(\boldsymbol{p}_1 \ldots, \boldsymbol{p}_n)$ that result from measurements $(\boldsymbol{\mathcal{A}}_1 \ldots, \boldsymbol{\mathcal{A}}_n)$ (counterfactually) applied to the same state ρ. A vast number of URs reported in the literature, including most entropic URs, take this form.

In fact, most of the entropic URs found so far fall under a much stronger restriction. As we mentioned in the previous section, they can all be constructed upon specific notions of joint uncertainty based on the "combined experiment" scenarios where either all the measurements are performed on independent, identically prepared quantum systems [as in Fig. 3.3], or a random choice is made to decide which of the several measurements to perform [as in Fig. 3.4]. All entropic relations based on joint

uncertainty quantifiers of the form $f(\boldsymbol{p}) + f(\boldsymbol{q})$, where f is an entropy function, fall under this category.

Going beyond these operational notions and using our general definition of joint uncertainty enables us to construct new URs, with the following "recipe":

1. Find a quantifier of the joint uncertainty (under a restricted class of symmetries, if applicable) of the desired number n of distributions, based on Def. 3.2.

2. For the given n measurements, find a lower bound on the n-joint uncertainty of the outcome distributions of the measurements applied to quantum states, like the c in (4.3). This bound leads to an assertion of the form (4.3), i.e. an uncertainty relation.

4.2.1 A New Quantum Uncertainty Relation

To demonstrate the utility of the aforementioned recipe, we derive an uncertainty relation for two projective measurements on pure state preparations of a 2-level system, using the following joint uncertainty quantifier constructed using Def. 3.2: $\mathcal{J}_2(\boldsymbol{p}, \boldsymbol{q}) = 1 - \boldsymbol{p}^{\downarrow} \cdot \boldsymbol{q}^{\downarrow}$. Here $(\cdot)$ denotes the usual dot product, and $\boldsymbol{p}^{\downarrow}$ is the rearraigned version of $\boldsymbol{p}$ in a non-increasing order. Note that this quantifier of joint uncertainty is *faithful* in the sense that it is zero if and only if both vectors $\boldsymbol{p}$ and $\boldsymbol{q}$ are completely certain.

Theorem 4.1. *Given a pair of projective measurements* $\boldsymbol{\mathcal{A}} = \{|a_1\rangle, |a_2\rangle\}$ *and* $\boldsymbol{\mathcal{B}} = \{|b_1\rangle, |b_2\rangle\}$*, for any quantum state* $|\psi\rangle$*,*

$$1 - \boldsymbol{p}^{\downarrow}(\psi) \cdot \boldsymbol{q}^{\downarrow}(\psi) \geq \frac{1}{2}(1 - \eta^2), \quad \forall \psi \tag{4.5}$$

where $\boldsymbol{p}$ and $\boldsymbol{q}$ are the outcome distributions of the projective measurements $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$; and $\eta \equiv \max_{i,j} |\langle a_i|b_j\rangle|$. Also, $\boldsymbol{p}^{\downarrow}$ denotes the sorted version of $\boldsymbol{p}$ in a non-increasing order.

Proof. Given a two-level quantum system (a qubit) in a pure state $|\psi\rangle$, consider two rank-1 projective measurements $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$, respectively defined by the orthonormal bases $\{|a_1\rangle, |a_2\rangle\}$ and $\{|b_1\rangle, |b_2\rangle\}$. When $\boldsymbol{\mathcal{A}}$ is applied to $|\psi\rangle$, the resulting outcome distribution $\boldsymbol{p}(\psi)$ has the components $p \equiv p_1(\psi) = |\langle a_1|\psi\rangle|^2$ and $p_2(\psi) = 1 - p$; similarly, we denote the distribution of the outcomes of $\boldsymbol{\mathcal{B}}$, $\boldsymbol{q}(\psi) \equiv (q, 1-q)$. We shall now find a lower bound on the minimum joint uncertainty of $(\boldsymbol{p}(\psi), \boldsymbol{q}(\psi))$, over all pure states $|\psi\rangle$, under the quantifier

$$\mathcal{J}_2(\boldsymbol{p}, \boldsymbol{q}) = 1 - \boldsymbol{p}^{\downarrow} \cdot \boldsymbol{q}^{\downarrow}. \tag{4.6}$$

Note that $\mathcal{J}_2$ is a valid quantifier of joint uncertainty as it satisfies the constraints in Definition 3.2. We can partition the set of all pure states into two subsets $\mathcal{S}_1$ and $\mathcal{S}_2$ given by

$$\mathcal{S}_1 = \{|\psi\rangle | (p \geq 0.5, q \geq 0.5) \text{ or } (p < 0.5, q < 0.5)\};$$
$$\mathcal{S}_2 = \{|\psi\rangle | (p \geq 0.5, q < 0.5) \text{ or } (p < 0.5, q \geq 0.5)\},$$

where p and q are understood to be ψ-dependent. The function $\mathcal{J}_2$ can be defined piecewise using this partition as

$$\mathcal{J}_2(\psi) = \begin{cases} p + q - 2pq, & |\psi\rangle \in \mathcal{S}_1; \\ 1 - p - q + 2pq, & |\psi\rangle \in \mathcal{S}_2. \end{cases} \tag{4.7}$$

Modulo a global phase, $|\psi\rangle$ can be parametrized as

$$|\psi\rangle = \cos\alpha |b_1\rangle + \mathrm{e}^{\mathrm{i}\varphi} \sin\alpha |b_2\rangle, \tag{4.8}$$

with $\alpha \in [0, \pi/2]$ and $\varphi \in [0, 2\pi)$. Appropriate global phases can be added to the measurement basis vectors so that

$$\begin{aligned} q &= \cos^2 \alpha; \\ p &= \left|\cos\alpha\cos\beta + \sin\alpha\sin\beta \mathrm{e}^{\mathrm{i}\varphi}\right|^2, \end{aligned}$$

where $\cos(\beta) \equiv |\langle a_1|b_1\rangle|$. The minimization of the function $\mathcal{J}_2$ using its piecewise definition (4.7) can be done by separately minimizing over $\mathcal{S}_1$ and $\mathcal{S}_2$ and then finding the smaller of the two minimums. Let us first consider $\mathcal{S}_1$. Now, it can be verified that the φ dependence of $\mathcal{J}_2$ is through a term of the form $f(\alpha, \beta)\sin^2(\varphi/2)$, so that the minimization can be carried out first over φ alone, and then over all α. In the cases where $f(\alpha, \beta) > 0$, the minimum over φ is achieved when $\sin^2(\varphi/2) = 1$; if $f(\alpha, \beta) < 0$, the minimum occurs when $\sin^2(\varphi/2) = 0$. In either case, the minimum over φ, as a function of α, takes the form

$$\begin{aligned} \min_{\varphi} \mathcal{J}_2(\psi) &\equiv J(\alpha) \\ &= \cos^2\alpha + \cos^2(\beta \pm \alpha) - 2\cos^2\alpha\cos^2(\beta \pm \alpha). \end{aligned}$$

Since $J(\alpha)$ is even in α, the subsequent minimization over α leads to the same value regardless of whether the positive or negative sign is used in the $\pm$ above. One can check that the minimum is attained at $\alpha = \beta/2$, yielding

$$\min_{|\psi\rangle \in \mathcal{S}_1} \mathcal{J}_2 = \frac{1}{2}\sin^2\beta. \tag{4.9}$$

Using similar arguments, we can determine the minimum over the other partition $\mathcal{S}_2$:

$$\min_{|\psi\rangle \in \mathcal{S}_2} \mathcal{J}_2 = \frac{1}{2}\cos^2\beta. \tag{4.10}$$

Note that without loss of generality we can take $0 \leq \beta \leq \pi/2$. Comparing the two local minimums, we can express the global minimum succinctly as

$$\min_{|\psi\rangle} \mathcal{J}_2(\boldsymbol{p}, \boldsymbol{q}) = \frac{1}{2}(1 - \eta^2), \tag{4.11}$$

where $\eta = \max_{i,j} |\langle a_i | b_j \rangle|$. □

Since the quantifier $\mathcal{J}_2$ cannot be interpreted based on the two combined-experiment scenarios under which most existing entropic URs fall, or indeed based on any single-experiment scenario, the above UR is non-trivially different from all previous ones. More generally, for d-dimensional $\boldsymbol{p}$ and $\boldsymbol{q}$, any joint uncertainty quantifier constructed as a Schur-concave function of the vector $(p_1^{\downarrow} q_1^{\downarrow}, p_2^{\downarrow} q_2^{\downarrow} \ldots, p_d^{\downarrow} q_d^{\downarrow})$ is a valid joint uncertainty quantifier of $(\boldsymbol{p}, \boldsymbol{q})$. So is any Schur-concave function of vectors of dimension $k < d$ constructed with the components $(p_1^{\downarrow} + q_1^{\downarrow}, p_2^{\downarrow} + q_2^{\downarrow} \ldots, p_k^{\downarrow} + q_k^{\downarrow})$. These are just a couple of examples that we contrived for illustration, suggesting that a rich variety of URs could be obtained by allowing joint uncertainty quantifiers that don't yield themselves to interpetation as the outcome uncertainty of any single experiment.

4.3 Fine-Grained Uncertainty Relations

Is it possible to find a *universal* relation that is independent of any quantifier of joint uncertainty, such that it generalizes all possible uncertainty relations with the generic form of (4.4), valid for any arbitrary quantifier $\mathcal{J}$? In this section we argue that such a universal relation does not exist. However, under some circumstances we are able to prove a large class of *quantifier-free* uncertainty relations.

Assume we intend to characterize the joint uncertainty of a given collection of quantum measurements; for convenience, suppose it is a pair of measurements $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}})$. We could construct various uncertainty relations using the aforementioned recipe,

with the given pair $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}})$ and different joint uncertainty quantifiers $\mathcal{J}$. Every uncertainty relation is stated in terms of a lower bound like the c of (4.3), which in turn depends on $\mathcal{J}$. In general, for a given $\mathcal{J}$ it might be hard to compute such a bound; it would be convenient if there were a shortcut.

For example, perhaps we could find a fixed pair $(\boldsymbol{s}, \boldsymbol{t})$ of distribution vectors (that only depend on $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}})$), and some mathematically defined binary relation $\preccurlyeq$ such that

$$(\boldsymbol{p}(\rho), \boldsymbol{q}(\rho)) \preccurlyeq (\boldsymbol{s}, \boldsymbol{t}) \quad \forall \rho$$

means

$$\mathcal{J}\left(\boldsymbol{p}(\rho), \boldsymbol{q}(\rho)\right) \geq \mathcal{J}(\boldsymbol{s}, \boldsymbol{t}) \quad \forall \rho, \mathcal{J}. \tag{4.12}$$

If there *were* such a pair, then for any given $\mathcal{J}_0$ we would merely have to compute $\mathcal{J}_0(\boldsymbol{s}, \boldsymbol{t})$, immediately yielding a bound. In this sense, finding such a pair would amount to finding a plethora of uncertainty relations; therefore, such a pair can be said to constitute a *quantifier-free uncertainty relation* for the pair $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}})$ [15, 16].

Of course, there is always a trivial possibility of such a pair: Just choose $\boldsymbol{s}_0$ and $\boldsymbol{t}_0$ such that $\boldsymbol{s}_0 \succ \boldsymbol{p}(\rho)$ and $\boldsymbol{t}_0 \succ \boldsymbol{q}(\rho)$ for all ρ. Such a $(\boldsymbol{s}_0, \boldsymbol{t}_0)$ would be unhelpful in that it would not impose *joint* restrictions on $(\boldsymbol{p}, \boldsymbol{q})$. What we would really like is a nontrivial pair $(\boldsymbol{s}, \boldsymbol{t})$ that does better. As it turns out, such a nontrivial pair satisfying (4.12) never exists for any given $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}})$, because the clause "$\forall \mathcal{J}$" in (4.12) includes all single-uncertainty quantifiers of $\boldsymbol{p}$ and $\boldsymbol{q}$ alone, leading necessarily to the trivial choice $(\boldsymbol{s}_0, \boldsymbol{t}_0)$. In particular, if the POVM associated with the measurement $\boldsymbol{\mathcal{A}}$ contains any rank-1 operators, then the condition $\boldsymbol{s} \succ \boldsymbol{p}(\rho)$ for all ρ is possible only if $\boldsymbol{s} = \boldsymbol{e} \equiv (1, 0, \cdots, 0)$; likewise for $\boldsymbol{\mathcal{B}}$.

In order to salvage the possibility of nontrivial quantifier-free uncertainty relations, we can relax the condition "$\forall \mathcal{J}$", and instead require the inequality in (4.12) to only

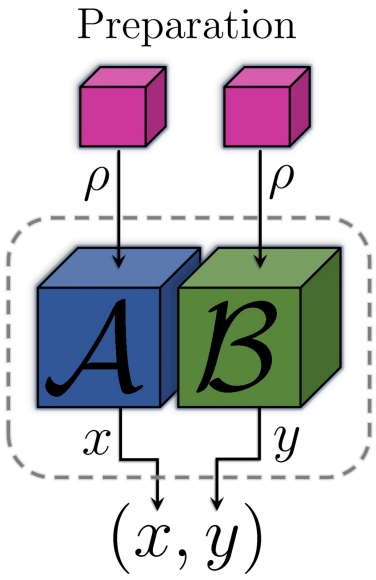


Figure 4.1: Type (a) combination of quantum measurements.

hold for some restricted class of $\mathcal{J}$'s. Sometimes such restricted classes of quantifiers do admit nontrivial $(\boldsymbol{s}, \boldsymbol{t})$ pairs. Here we find two such restricted classes.

Interestingly, these classes are related to the two examples we considered in Section 3.4. The first scenario considers the experiment where *both* $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$ are carried out independently of each other, on the *same* state ρ. Since measurements in general destroy the state of a quantum system, independently carrying out both $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$ requires us to first prepare two independent (i.e. uncorrelated) copies of the state ρ (Fig. 4.1). These copies, represented by the tensor product $\rho \otimes \rho$, are then subjected to the measurements $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$, and the outcome of this joint experiment is a random variable $Z \equiv (X, Y)$.

We can then restrict to those joint uncertainty quantifiers $\mathcal{J}$ that are obtained as quantifiers of the *single-variable* uncertainty of the variable Z. Therefore the quantifiers in our restricted class are of the form

$$\mathcal{J}\left(\boldsymbol{p}(\rho), \boldsymbol{q}(\rho)\right) = \mathcal{U}\left(\boldsymbol{p}(\rho) \otimes \boldsymbol{q}(\rho)\right), \tag{4.13}$$

where $\mathcal{U}$ is a single-variable uncertainty quantifier. It was shown [15, 16] that for any given projective measurements $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}})$ there exists a distribution vector $\boldsymbol{s} \equiv \boldsymbol{s}(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}})$ such that

$$\boldsymbol{p}(\rho) \otimes \boldsymbol{q}(\rho) \prec \boldsymbol{s}, \quad \forall \rho. \tag{4.14}$$

This uncertainty relation is a fine-grained uncertainty relation, since it is independent of the choice of the function used to quantify the uncertainty.

By applying any quantifier of uncertainty (any Schur-concave function $\mathcal{U}$) to the quantifier-free uncertainty relation (4.14) we will have

$$\mathcal{U}\left(\boldsymbol{p}(\rho) \otimes \boldsymbol{q}(\rho)\right) \geq \mathcal{U}(\boldsymbol{s}) = \text{const.}, \quad \forall \rho. \tag{4.15}$$

Note that for Rényi entropies (and in general any Schur-concave function that is additive under direct product, i.e. $\mathcal{U}(\boldsymbol{p} \otimes \boldsymbol{q}) = \mathcal{U}(\boldsymbol{p}) + \mathcal{U}(\boldsymbol{q})$) equation (4.15) becomes an entropic uncertainty relation (EUR) of the following form:

$$H_\alpha(\boldsymbol{p}(\rho)) + H_\alpha(\boldsymbol{q}(\rho)) \geq H_\alpha(\boldsymbol{s}) = \text{const.}, \quad \forall \rho. \tag{4.16}$$

Similarly, following Example 2 of Section 3.4, we can consider another combination wherein we first pick, at random, only one of the two projective measurements $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$, and then perform it on a pure state $\rho = |\psi\rangle\langle\psi|$. This is the same scenario considered in [7]. For this case, we can find a nontrivial $(\boldsymbol{s}, \boldsymbol{t})$.

Theorem 4.2. *For a given pair of projective measurements, there exists a non-trivial pair of distributions $(\boldsymbol{s}, \boldsymbol{t})$, independent of $|\psi\rangle$, such that, for any $0 \leq r \leq 1$ and any quantum state $|\psi\rangle$*

$$\mathcal{U}(r\boldsymbol{p}(\psi) \oplus (1-r)\boldsymbol{q}(\psi)) \geq \mathcal{U}(r\boldsymbol{s} \oplus (1-r)\boldsymbol{t}), \quad \forall\ |\psi\rangle \tag{4.17}$$

where $\boldsymbol{p}$ and $\boldsymbol{q}$ are probability vectors respectively associated with the two measurements performed on $|\psi\rangle$; and $\mathcal{U}$ is a Schur-concave function.

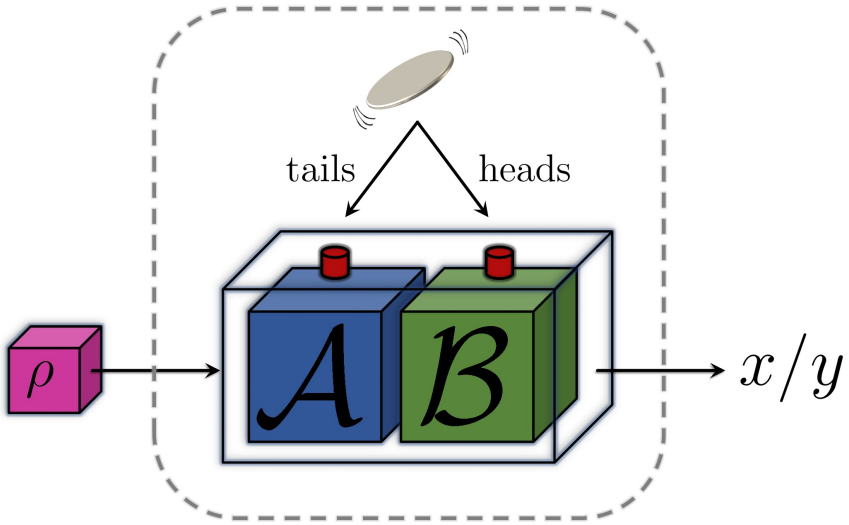


Figure 4.2: Type (b) combination of quantum measurements.

Proof. Let $\boldsymbol{\mathcal{A}} \equiv \{\Lambda_x = |a_x\rangle\langle a_x|\}$ and $\boldsymbol{\mathcal{B}} \equiv \{\Gamma_y = |b_y\rangle\langle b_y|\}$ be two rank-1 projective measurements on a d-dimensional quantum system. Consider a particular combination of these measurements, shown in Fig. 4.2: We first randomly choose one of $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$, as per the distribution $(r, 1-r)$ (without loss of generality we assume $r \geq 0.5$); we then perform the chosen measurement. Let $p_x = |\langle a_x|\psi\rangle|^2$ and $q_y = |\langle b_y|\psi\rangle|^2$ denote the output probability distributions of measurements. Thus, the output probability vector of this combined setup will have the following form

$$\boldsymbol{h}(\psi) = (r\boldsymbol{p}(\psi)) \oplus ((1-r)\boldsymbol{q}(\psi)), \tag{4.18}$$

understood as the vector obtained by concatenating the components of the two vectors $r\boldsymbol{p}$ and $(1-r)\boldsymbol{q}$.

If there exist a probability vector $\boldsymbol{\omega}(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}}) \equiv r\boldsymbol{s} \oplus (1-r)\boldsymbol{t}$ such that

$$\boldsymbol{h}(\psi) \prec \boldsymbol{\omega} \quad \forall\ |\psi\rangle, \tag{4.19}$$

we can immediately have

$$\mathcal{J}\left(\boldsymbol{p}(\psi), \boldsymbol{q}(\psi)\right) \geq \mathcal{J}(\boldsymbol{s}, \boldsymbol{t}) \quad \forall\ |\psi\rangle, \tag{4.20}$$

for all the Schur-concave functions $\mathcal{J}(\boldsymbol{p},\boldsymbol{q}) \equiv \mathcal{U}(r\boldsymbol{p} \oplus (1-r)\boldsymbol{q})$. Hence, 4.19 forms a quantifier-free quantum uncertainty relation.

We shall now find such an $\boldsymbol{\omega}$. Let $\boldsymbol{h}^{\downarrow}$ be the vector $\boldsymbol{h}$ with its components arranged in a non-increasing order. Then define

$$\Omega_k \equiv \max_{|\psi\rangle} \sum_{j=1}^{k} \left(\boldsymbol{h}^{\downarrow}(\psi)\right)_j, \tag{4.21}$$

where $1 \leq k \leq 2d$. According to theorem 2.2, probability vector $\boldsymbol{\omega}$ constructed as

$$\boldsymbol{\omega} = (\Omega_1, \Omega_2 - \Omega_1, \cdots, \Omega_{2d} - \Omega_{2d-1}), \tag{4.22}$$

majorizes $\boldsymbol{h}(\psi)$ for all $|\psi\rangle$.

It is straightforward to prove that $\Omega_1 = r$, and that $\Omega_k = 1$ for all $d+1 \leq k \leq 2d$ (take $|\psi\rangle = |a_{x_0}\rangle$ and $|\psi\rangle = |b_{y_0}\rangle$ respectively, where x_0 and y_0 are specific instances of x and y). Also, in the following lemma we find Ω_k, for $2 \leq k \leq d$, using minor modifications of the lemma in the appendix of Ref. [16]. Thus, in a simplified form:

$$\boldsymbol{\omega} = (r, \Omega_2 - r, \cdots, 1 - \Omega_d, 0, \cdots, 0). \tag{4.23}$$

As the order of components is not important for the majorization relation, we can rewrite the majorizing probability vector $\boldsymbol{\omega}$ as

$$\boldsymbol{\omega} = r\boldsymbol{e} \oplus (1-r)\boldsymbol{t}_r, \tag{4.24}$$

where $\boldsymbol{e} = (1, 0, 0, \cdots)$ and the probability vector $\boldsymbol{t}_r$ is given by

$$\boldsymbol{t}_r(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}}) = \frac{1}{1-r}\left(\Omega_2 - r, \Omega_3 - \Omega_2, \cdots\right). \tag{4.25}$$

Clearly, the pair $(\boldsymbol{e}, \boldsymbol{t}_r)$ is the desired $(\boldsymbol{s}, \boldsymbol{t})$ of theorem 4.2. □

Lemma 4.1. *Given two projective measurements $\boldsymbol{\mathcal{A}} \equiv \{|a_x\rangle\langle a_x|\}$ and $\boldsymbol{\mathcal{B}} \equiv \{|b_y\rangle\langle b_y|\}$, define the matrix $G = (g_{xy})$ with $g_{xy} = \langle a_x|b_y\rangle$. Also define the set $\mathcal{S}(G,k)$ of all submatrices of G, the numbers of whose rows and columns together add up to k (i.e. all*

submatrices with 1 row and $(k-1)$ columns, all submatrices with 2 rows and $(k-2)$ columns, etc.). Let $p_x = |\langle a_x|\psi\rangle|^2$ and $q_y = |\langle b_y|\psi\rangle|^2$ denote the output probability distributions of measurements. According to (4.18) *let $\boldsymbol{h} = r\boldsymbol{p} \oplus (1-r)\boldsymbol{q}$; then, $\Omega_k \equiv \max_{|\psi\rangle} \sum_{j=1}^{k} \left(\boldsymbol{h}^{\downarrow}(\psi)\right)_j$, can be obtained via the following formula*

$$\Omega_k = r + \sqrt{r(1-r)} \left(\max_{V \in \mathcal{S}(G,k)} \sigma_1(V) \right), \tag{4.26}$$

where, $\sigma_1(V)$ is the leading singular value of the rectangular matrix V (Note that singular value decomposition is invariant over permutations of rows and columns of V).

Proof. Assume that we know the first k elements of $\boldsymbol{h}^{\downarrow}(\psi)$: it has k_1 p_x's and k_2 q_y's with $k_1 + k_2 = k$. Then:

$$\Omega_k = \max_{|\psi\rangle} \sum_{j=1}^{k_1} r\tilde{p}_j + \sum_{j=1}^{k_2} (1-r)\tilde{q}_j \tag{4.27}$$

$$= \max_{|\psi\rangle} \sum_{j=1}^{k_1} r|\langle \tilde{a}_j|\psi\rangle|^2 + \sum_{j=1}^{k_2} (1-r)|\langle \tilde{b}_j|\psi\rangle|^2, \tag{4.28}$$

where $\tilde{\boldsymbol{p}}$ and $\tilde{\boldsymbol{q}}$ are some permuted versions of $\boldsymbol{p}$ and $\boldsymbol{q}$ (not necessarily the same permutation). Let $|\hat{a}_j\rangle = \sqrt{r}|\tilde{a}_j\rangle$ and $|\hat{b}_j\rangle = \sqrt{1-r}|\tilde{b}_j\rangle$, then we can rewrite Ω_k as

$$\Omega_k = \max_{|\psi\rangle} \sum_{j=1}^{k_1} |\langle \hat{a}_j|\psi\rangle|^2 + \sum_{j=1}^{k_2} |\langle \hat{b}_j|\psi\rangle|^2 \tag{4.29}$$

$$= \max_{|\psi\rangle} ||C|\psi\rangle||^2 = \sigma_1^2(C) = \lambda_1(CC^{\dagger}), \tag{4.30}$$

where $\sigma_1(\cdot)$ and $\lambda_1(\cdot)$ denote the leading singular value and leading eigenvalue, re-

spectively; and the matrix C is defined as:

$$C = \begin{bmatrix} \langle \hat{a}_1| \\ \vdots \\ \langle \hat{a}_{k_1}| \\ \langle \hat{b}_1| \\ \vdots \\ \langle \hat{b}_{k_2}| \end{bmatrix} = \begin{bmatrix} \sqrt{r}\langle \tilde{a}_1| \\ \vdots \\ \sqrt{r}\langle \tilde{a}_{k_1}| \\ \sqrt{1-r}\langle \tilde{b}_1| \\ \vdots \\ \sqrt{1-r}\langle \tilde{b}_{k_2}| \end{bmatrix}. \tag{4.31}$$

Note that

$$C^\dagger = \left[\sqrt{r}|\tilde{a}_1\rangle \ \cdots \ \sqrt{r}|\tilde{a}_{k_1}\rangle \ \sqrt{1-r}|\tilde{b}_1\rangle \ \cdots \ \sqrt{1-r}|\tilde{b}_{k_2}\rangle \right], \tag{4.32}$$

and consequently

$$CC^\dagger = \begin{bmatrix} r\mathbb{1}_{k_1} & \sqrt{r(1-r)}V \\ \sqrt{r(1-r)}V^\dagger & r\mathbb{1}_{k_2} \end{bmatrix}, \tag{4.33}$$

with $V = (v_{ij})$, where $v_{ij} = \langle \tilde{a}_i|\tilde{b}_j\rangle$. Now according to the Jordan's definition of singular values, we can simplify $\lambda_1(CC^\dagger)$ as

$$\lambda_1(CC^\dagger) = r + \sqrt{r(1-r)}\lambda_1 \begin{pmatrix} 0 & V \\ V^\dagger & 0 \end{pmatrix} \tag{4.34}$$

$$= r + \sqrt{r(1-r)}\sigma_1(V). \tag{4.35}$$

As we don't know the optimizing combination of p_x's and q_y's we need to maximize over all possible V matrices with k_1 rows and k_2 columns such that $k_1 + k_2 = k$. This optimization leads to (4.26) which completes the proof. □

In conclusion, we constructed a new fine-grained uncertainty relation:

$$(r\boldsymbol{p}(\psi)) \oplus ((1-r)\boldsymbol{q}(\psi)) \prec r\boldsymbol{e} \oplus (1-r)\boldsymbol{t}_r, \quad \forall\ |\psi\rangle \tag{4.36}$$

where $\boldsymbol{p}$ and $\boldsymbol{q}$ are output probability distributions associated with quantum projective measurements $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$, respectively; and $\boldsymbol{t}_r$ is defined in (4.25), where r is any arbitrary real number $0.5 \leq r \leq 1$.

Now by applying any Schur-concave function $\mathcal{U}$, the fine-grained uncertainty relation (4.36) immediately leads to the following uncertainty relation based on the uncertainty quantifier $\mathcal{U}$:

$$\mathcal{U}\left((r\boldsymbol{p}(\psi)) \oplus ((1-r)\boldsymbol{q}(\psi))\right) \prec \mathcal{U}\left(r\boldsymbol{e} \oplus (1-r)\boldsymbol{t}_r\right), \quad \forall\ |\psi\rangle. \tag{4.37}$$

For example, for the case of Shannon entropy ($\mathcal{U} = H_1$), and $r = 0.5$, our uncertainty relation simply generates the following entropic uncertainty relation:

$$H_1(\boldsymbol{p}(\psi)) + H_1(\boldsymbol{q}(\psi)) \geq H_1(\boldsymbol{t}_{0.5}) = \text{const.}, \quad \forall\ |\psi\rangle \tag{4.38}$$

where $H_1(\boldsymbol{t}_{0.5})$ is a positive constant independent of $|\psi\rangle$.

4.4 Conclusion

We took our axiomatic approach discussed in chapter 3 and used it to generalize a large class of uncertainty relations into a single framework. Our framework also provides a recipe to discover new uncertainty relations. We gave one such example of new quantum uncertainty relations that was not discussed in the literature.

We then studied fine-grained (or quantifier-free) quantum uncertainty relations. We showed that no universal quantifier-free relation exists that can generalize all uncertainty relations. Nonetheless, we proved a fine-grained majorization uncertainty relation that generalizes a restricted, but a large class of entropic uncertainty relations. These majorization relations are fine-grained in the sense that they provide a direct constraint on the probability distribution vectors that are associated with the two quantum measurement outcomes.

Chapter 5

Conclusion and Open Problems

In this work, we indicated that uncertainty is a property of a physical system that cannot be reduced by mere relabeling or by random relabeling. One can further restrict allowed relabelings to a particular symmetry group depending on the physical system.

We based our analysis on an information-theoretic study of the mechanisms of uncertainty increase: 1) randomly-chosen symmetry transformations; and 2) classical processing via channels followed by recovery. Corresponding to these, we identified two classes of doubly stochastic matrices, $\mathcal{D}^{\text{sym}}$ and $\mathcal{D}^{\text{rec}}$. Uncertainty quantifiers, in the strictest sense, must be monotonically non-decreasing under the action of these two classes.

We then took a similar information-theoretic approach to the concept of joint uncertainty of several variables, resulting in the principle that the most basic features of joint uncertainty quantifiers must not depend on specific operational combinations of the variables. We then considered quantum uncertainty relations (URs) of the preparational uncertainty type, where past works have always assumed specific operational combinations. Applying our new notion of joint uncertainty not only resulted in a unified understanding of a large class of URs, but also opened up the possibility of deriving a new class of preparational URs, namely identities that are mathematically valid for any preparation, but cannot be interpreted based on any single experimental scenario. To illustrate, we constructed a class of joint uncertainty quantifiers with this property, and proved a new quantum uncertainty relation based on one of these

quantifiers as an example.

Finally, we found that so-called universal uncertainty relations cannot be found over all possible quantifiers of joint uncertainty. We connected fine-grained uncertainty relations found in past works (see [15, 16, 43, 44]) with specific operational interpretations of joint uncertainty. Considering one such operational scenario we also derived a new class of fine-grained uncertainty relations.

In addition, we noted the limitations of all fine-grained uncertainty relations. They are powerful tools inasmuch as they generate a variety of uncertainty relations, but we must bear in mind that the bounds they yield on a specific joint uncertainty quantifier may not be tight.

5.1 Future Work

There is much to be understood about the classes $\mathcal{D}^{\text{rec}}$ and $\mathcal{D}^{\text{sym}}$ of doubly stochastic matrices. First open question here is to find a rigorous characterization of $\mathcal{D}^{\text{rec}}$. Second question is to find the definition of the hierarchy relations among real vectors induced by the action of doubly stochastic matrices in the union of these two classes. In other words, it is interesting to find the necessary and sufficient conditions on the elements of the two given real vectors $\boldsymbol{p}$ and $\boldsymbol{q}$, if $\boldsymbol{q} = D\boldsymbol{p}$ with D belonging to the union of $\mathcal{D}^{\text{rec}}$ and $\mathcal{D}^{\text{sym}}$. One other interesting question would be to find quantifiers of uncertainty that are not Schur-concave but are concave under the action of the matrices in the union of $\mathcal{D}^{\text{rec}}$ and $\mathcal{D}^{\text{sym}}$.

We introduced and discussed two classes of quantifier-free uncertainty relations in section 4.3. It might be possible to unify the spirit of the two classes of the majorization URs mentioned in the previous chapter into a larger class, by including all quantifiers of joint uncertainty that are *symmetric* in the two (or more) distributions:

$$\mathcal{J}(\boldsymbol{p},\boldsymbol{q}) = \mathcal{J}(\boldsymbol{q},\boldsymbol{p}). \tag{5.1}$$

This requirement avoids the case of trivial relations resulting from the requirement $(\boldsymbol{s},\boldsymbol{t}) \gg (\boldsymbol{p},\boldsymbol{q})$, but we leave it open whether a nontrivial $(\boldsymbol{s},\boldsymbol{t})$ can be found. Another way of unifiying several classes of quantifier-free relations, each with its respective $(\boldsymbol{s}_i,\boldsymbol{t}_i)$, is by bounding any quantifier $\mathcal{J}$ as follows:

$$\mathcal{J}\left(\boldsymbol{p}(\rho),\boldsymbol{q}(\rho)\right) \geq \min_{j\in\{1,\ldots,m\}} \mathcal{J}(\boldsymbol{s}_j,\boldsymbol{t}_j) \quad \forall \rho, \mathcal{J}. \tag{5.2}$$

An interesting open problem is whether there exists a finite integer m such that minimizing over all $j \leq m$ provides a nontrivial bound for all nontrivial joint uncertainty quantifiers.

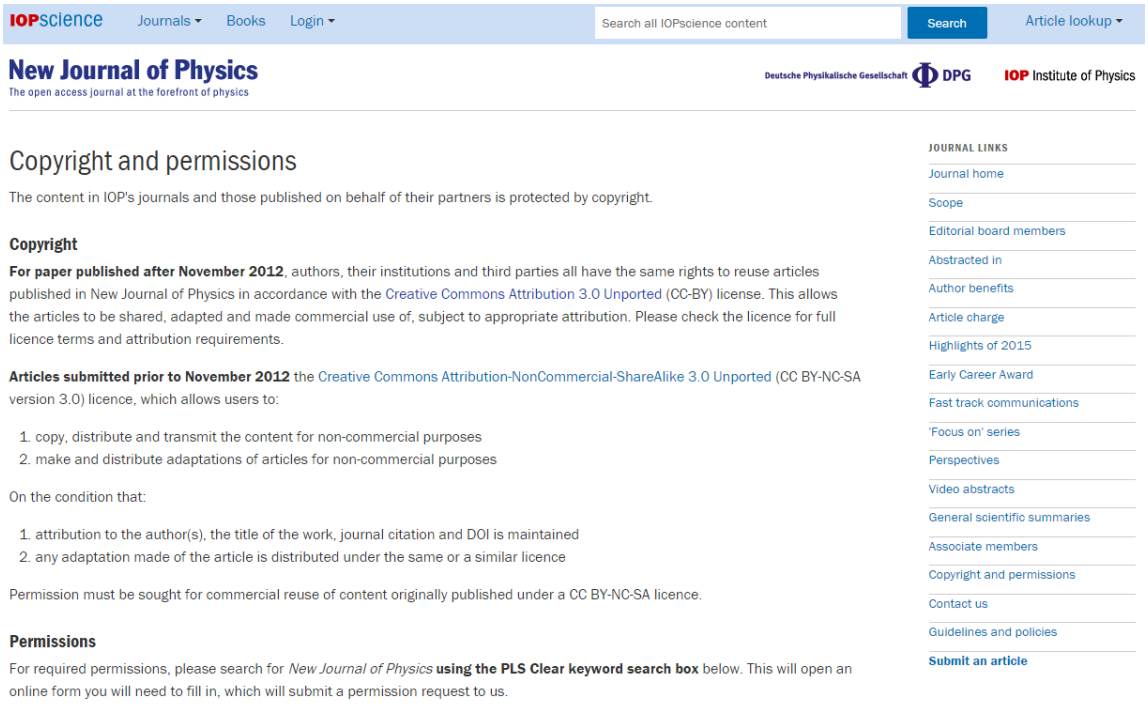
Moreover, a more complete characterization of uncertainty on infinite-dimensional systems is another challenging future project. This could impact applications of squeezed states, which are ubiquitous in quantum information processing with continuous variables.

One significance of our work is our approach to the notions of uncertainty and joint uncertainty. It would be very intriguing to find more applications of this approach in other areas of research, such as classical information theory, cryptography, etc.

Appendix A

Copyright and Permissions

This thesis uses the material from Ref. [3] under the Creative Commons Attribution 3.0 Unported (CC-BY) license. For more documentation and information regarding the copyright and permissions of the New Journal of Physics please see figure below or go to `http://iopscience.iop.org/journal/1367-2630/page/Copyright%20and%20permissions`.



IOPscience Journals Books Login Search all IOPscience content Search Article lookup

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG IOP Institute of Physics

Copyright and permissions

The content in IOP's journals and those published on behalf of their partners is protected by copyright.

Copyright

For paper published after November 2012, authors, their institutions and third parties all have the same rights to reuse articles published in New Journal of Physics in accordance with the Creative Commons Attribution 3.0 Unported (CC-BY) license. This allows the articles to be shared, adapted and made commercial use of, subject to appropriate attribution. Please check the licence for full licence terms and attribution requirements.

Articles submitted prior to November 2012 the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported (CC BY-NC-SA version 3.0) licence, which allows users to:

1. copy, distribute and transmit the content for non-commercial purposes
2. make and distribute adaptations of articles for non-commercial purposes

On the condition that:

1. attribution to the author(s), the title of the work, journal citation and DOI is maintained
2. any adaptation made of the article is distributed under the same or a similar licence

Permission must be sought for commercial reuse of content originally published under a CC BY-NC-SA licence.

Permissions

For required permissions, please search for *New Journal of Physics* **using the PLS Clear keyword search box** below. This will open an online form you will need to fill in, which will submit a permission request to us.

JOURNAL LINKS
Journal home
Scope
Editorial board members
Abstracted in
Author benefits
Article charge
Highlights of 2015
Early Career Award
Fast track communications
'Focus on' series
Perspectives
Video abstracts
General scientific summaries
Associate members
Copyright and permissions
Contact us
Guidelines and policies
Submit an article

Figure A.1: Copyright and reuse permissions of the New Journal of Physics.

Bibliography

[1] W. Heisenberg, "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik," *Zeitschrift für Physik*, vol. 43, pp. 172–198, Mar. 1927.

[2] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[3] V. Narasimhachar, A. Poostindouz, and G. Gour, "Uncertainty, joint uncertainty, and the quantum uncertainty principle," *New Journal of Physics*, vol. 18, p. 033019, Mar. 2016.

[4] H. Maassen and J. Uffink, "Generalized entropic uncertainty relations," *Physical Review Letters*, vol. 60, pp. 1103–1106, Mar. 1988.

[5] O. Gühne, "Characterizing Entanglement via Uncertainty Relations," *Physical Review Letters*, vol. 92, p. 117903, Mar. 2004.

[6] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, "The uncertainty principle in the presence of quantum memory," *Nature Physics*, vol. 6, pp. 659–662, Jul. 2010.

[7] J. Oppenheim and S. Wehner, "The uncertainty principle determines the nonlocality of quantum mechanics.," *Science (New York, N.Y.)*, vol. 330, pp. 1072–4, Nov. 2010.

[8] S. Wehner and A. Winter, "Entropic uncertainty relations—a survey," *New Journal of Physics*, vol. 12, p. 025009, Feb. 2010.

[9] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, "Entropic Uncertainty Relations and their Applications," *arXiv:1511.04857 [quant-ph]*, pp. 1–64, Nov. 2015.

[10] Y. Xiao, N. Jing, S.-M. Fei, T. Li, X. Li-Jost, T. Ma, and Z.-X. Wang, "Strong entropic uncertainty relations for multiple measurements," *Phys. Rev. A*, vol. 93, p. 042125, Apr. 2016.

[11] F. Adabi, S. Salimi, and S. Haseli, "Tightening the entropic uncertainty bound in the presence of quantum memory," *Phys. Rev. A*, vol. 93, p. 062123, Jun. 2016.

[12] J. Zhang, Y. Zhang, and C.-s. Yu, "Entropic Uncertainty Relation and Information Exclusion Relation for multiple measurements in the presence of quantum memory," *Scientific Reports*, vol. 5, p. 11701, Jun. 2015.

[13] S. Liu, L.-Z. Mu, and H. Fan, "Entropic uncertainty relations for multiple measurements," *Phys. Rev. A*, vol. 91, p. 042133, Apr. 2015.

[14] M. H. Partovi, "Majorization formulation of uncertainty in quantum mechanics," *Physical Review A*, vol. 84, p. 052117, Nov. 2011.

[15] S. Friedland, V. Gheorghiu, and G. Gour, "Universal Uncertainty Relations," *Physical Review Letters*, vol. 111, p. 230401, Dec. 2013.

[16] Z. Puchała, Ł. Rudnicki, and K. Życzkowski, "Majorization entropic uncertainty relations," *Journal of Physics A: Mathematical and Theoretical*, vol. 46, p. 272002, Jun. 2013.

[17] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications*. Springer Series in Statistics, New York, NY: Springer New York, 2011.

[18] M. M. Wilde, *Quantum information theory*. Cambridge university press, 1st ed., 2013.

[19] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge university press, 10th anniv ed., 2010.

[20] G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities*. Cambridge university press, 1952.

[21] G. Birkhoff, "Three observations on linear algebra," *Revi. Univ. Nac. Tucuman, Ser. A*, vol. 5, pp. 147–151, 1946.

[22] H. Robertson, "The Uncertainty Principle," *Physical Review*, vol. 34, pp. 163–164, Jul. 1929.

[23] R. Bhatia, *Matrix Analysis*, vol. 169 of *Graduate Texts in Mathematics*. New York, NY: Springer New York, 1997.

[24] E. Kennard, "Zur quantenmechanik einfacher bewegungstypen," *Zeitschrift für Physik*, vol. 44, no. 4-5, pp. 326–352, 1927.

[25] I. I. Hirschman, "A Note on Entropy," *American Journal of Mathematics*, vol. 79, p. 152, Jan. 1957.

[26] W. Beckner, "Inequalities in Fourier Analysis," *The Annals of Mathematics*, vol. 102, p. 159, Jul. 1975.

[27] I. Białynicki-Birula and J. Mycielski, "Uncertainty relations for information entropy in wave mechanics," *Communications in Mathematical Physics*, vol. 44, pp. 129–132, Jun. 1975.

[28] W. Heisenberg, "Remarks on the origin of the relations of uncertainty," *The Uncertainty Principle and Foundation of Quantum Mechanics*, pp. 3–6, 1977.

[29] M. Ozawa, "Universally valid reformulation of the heisenberg uncertainty principle on noise and disturbance in measurement," *Physical Review A*, vol. 67, p. 042105, Apr. 2003.

[30] M. Ozawa, "Uncertainty relations for noise and disturbance in generalized quantum measurements," *Annals of Physics*, vol. 311, pp. 350–416, Jun. 2004.

[31] P. Busch, P. Lahti, and R. F. Werner, "Proof of heisenberg's error-disturbance relation," *Physical review letters*, vol. 111, p. 160405, Oct. 2013.

[32] D. Deutsch, "Uncertainty in Quantum Measurements," *Physical Review Letters*, vol. 50, pp. 631–633, Feb. 1983.

[33] I. D. Ivanovic, "An inequality for the sum of entropies of unbiased quantum measurements," *Journal of Physics A: Mathematical and General*, vol. 25, pp. L363–L364, Apr. 1992.

[34] J. Sánchez, "Entropic uncertainty and certainty relations for complementary observables," *Physics Letters A*, vol. 173, pp. 233–239, Feb. 1993.

[35] P. J. Coles and M. Piani, "Improved entropic uncertainty relations and information exclusion relations," *Physical Review A*, vol. 89, p. 022112, Feb. 2014.

[36] M. Berta, P. J. Coles, and S. Wehner, "Entanglement-assisted guessing of complementary measurement outcomes," *Phys. Rev. A*, vol. 90, p. 062127, Dec. 2014.

[37] J. Kaniewski, M. Tomamichel, and S. Wehner, "Entropic uncertainty from effective anticommutators," *Phys. Rev. A*, vol. 90, p. 012332, Jul. 2014.

[38] M. Tomamichel and R. Renner, "Uncertainty Relation for Smooth Entropies," *Physical Review Letters*, vol. 106, p. 110506, Mar. 2011.

[39] P. J. Coles, R. Colbeck, L. Yu, and M. Zwolak, "Uncertainty Relations from Simple Entropic Properties," *Physical Review Letters*, vol. 108, p. 210405, May 2012.

[40] M. Ballester and S. Wehner, "Entropic uncertainty relations and locking: Tight bounds for mutually unbiased bases," *Physical Review A*, vol. 75, p. 022319, Feb. 2007.

[41] S. Wu, S. Yu, and K. Mølmer, "Entropic uncertainty relation for mutually unbiased bases," *Physical Review A*, vol. 79, p. 022104, Feb. 2009.

[42] Y. Huang, "Entropic uncertainty relations in multidimensional position and momentum spaces," *Phys. Rev. A*, vol. 83, p. 052124, May 2011.

[43] Ł. Rudnicki, Z. Puchała, and K. Życzkowski, "Strong majorization entropic uncertainty relations," *Physical Review A*, vol. 89, p. 052115, May 2014.

[44] Ł. Rudnicki, "Majorization approach to entropic uncertainty relations for coarse-grained observables," *Physical Review A*, vol. 91, p. 032123, Mar. 2015.